

STIC-Biotech/ChemLib

118457

From: Whiteman, Brian
Sent: Thursday, April 01, 2004 1:00 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

09/845,416 4/30/01 Xiao Xiao
4/30/01

search nucleotides 900-3000 of SEQ ID NO: 2 against us patent and us patent application databases
search nucleotides 1000-3000 of SEQ ID NO: 6 against us patent and us patent application databases
search nucleotides 600-3000 of SEQ ID NO: 9 against us patent and us patent application databases
search nucleotides 1800-3100 of SEQ ID NO: 10 against us patent and us patent application databases
search nucleotides 900-3000 of SEQ ID NO: 12 against us patent and us patent application databases
search nucleotides 900-2500 of SEQ ID NO: 14 against us patent and us patent application databases

Thanks,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

CRF

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/1/04
Date Completed: 4/5/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: 6
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: both
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model
un on: April 4, 2004, 11:55:32 ; Search time 516.337 Seconds
(without alignments)
16463.377 Million cell updates/sec

itle: US-09-845-416-6_COPY_1000_3000
erfect score: 2001
equence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaac 2001

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

earched: 3373863 segs, 2124099041 residues

otal number of hits satisfying chosen parameters: 6747726

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB	ID	Description
1	2001	100.0	3999	6	AAD37234	Aad37234 Human dys
2	2001	100.0	4966	6	AAD37256	Aad37256 Adeno-ass
3	2001	100.0	4990	6	AAD37262	Aad37262 Adeno-ass
4	1709	85.4	3858	6	AAD37237	Aad37237 Human dys
5	1709	85.4	4825	6	AAD37257	Aad37257 Adeno-ass
6	1709	85.4	4848	6	AAD37263	Aad37263 Adeno-ass
7	1709	85.4	5060	6	AAD37264	Aad37264 Adeno-ass
8	1662.8	83.1	4182	6	AAD37230	Aad37230 Human dys
9	1662.8	83.1	5149	6	AAD37255	Aad37255 Adeno-ass
10	1283	64.1	5462	6	ABK81999	Abk81999 DNA encod
11	1182.6	59.1	5952	5	AAD06794	Aad06794 Human dys
12	1182.6	59.1	8689	6	ABK82000	Abk82000 DNA encod
13	1182.6	59.1	11058	6	AAD37229	Aad37229 Human dys
14	1182.6	59.1	11241	6	ABK82005	Abk82005 cDNA enco
15	1182.6	59.1	11443	6	ABK82002	Abk82002 DNA encod
16	1182.6	59.1	12923	1	AAN90338	Aan90338 Sequence
17	1182.6	59.1	13957	6	ABK81959	Abk81959 cDNA enco
18	1182.6	59.1	13957	6	ABT10904	Abt10904 Human bre
19	1182.6	59.1	13957	6	ABN95786	Abn95786 Gene #228
20	1182.6	59.1	13957	6	ABS69900	Abs69900 Human dys
21	1182.6	59.1	13977	6	ABS70403	Abs70403 Human bon
22	1182	59.1	2169	6	AAD37232	Aad37232 Human dys
23	1182	59.1	3531	6	AAD37238	Aad37238 Human dys

24	1182	59.1	4498	6	AAD37258	Aad37258 Adeno-ass
25	1180	59.0	5339	6	ABK81998	Abk81998 DNA encod
26	1013	50.6	3510	6	AAD37240	Aad37240 Human dys
27	1013	50.6	4476	6	AAD37259	Aad37259 Adeno-ass
28	1004	50.2	13815	2	AAV18885	Aav18885 Mus muscu
29	1004	50.2	13815	6	ABK81960	Abk81960 cDNA enco
30	1004	50.2	13815	6	ABI99799	Abi99799 Mouse isc
31	1004	50.2	19307	2	AAT27558	Aat27558 Shuttle v
32	835	41.7	1821	6	AAD37241	Aad37241 Human dys
33	727	35.3	5417	6	ABK81997	Abk81997 DNA encod
34	678.4	33.9	4402	3	AAZ48568	Aaz48568 A rod sho
35	678.4	33.9	4414	6	AAD37260	Aad37260 Adeno-ass
36	678	33.9	1991	6	AAD37231	Aad37231 Human dys
37	677	33.8	1667	6	AAD37235	Aad37235 Human dys
38	666.4	33.3	3446	6	AAD37242	Aad37242 Human dys
39	618.8	30.9	4075	3	AAZ48569	Aaz48569 A rod sho
40	614.2	30.7	4402	3	AAZ48567	Aaz48567 A rod sho
41	537.8	26.9	3275	1	AAN97129	Aan97129 Partial s
42	448	22.4	1434	6	AAD37243	Aad37243 Human dys
43	408.2	20.4	11096	6	ABK81962	Abk81962 cDNA enco
44	408	20.4	10705	7	ABT41896	Abt41896 Toxicity
45	397	19.8	6045	2	AAT74665	Aat74665 Utrophin

ALIGNMENTS

RESULT 1
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
(XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 6; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1059

2Y 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 1119

2Y 121 GATGTGGAAGTGGTGAAGAACCGATTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGAACCGATTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1179

2Y 181 GCCCATCAGGCCGGGTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACACAGGA 240
Db 1180 GCCCATCAGGCCGGGTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACACAGGA 1239

2Y 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCCTAAATTTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCCTAAATTTCAAGA 1299

2Y 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359

2Y 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGCTAAACAAACAGAAAGAAAGA 420
Db 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGCTAAACAAACAGAAAGAAAGA 1419

2Y 421 ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 480
Db 1420 ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 1479

2Y 481 CAACAACATTAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540
Db 1480 CAACAACATTAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 1539

2Y 541 ACTCACATGGTGGTAGTTGATGNAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTAGTTGATGNAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599

2Y 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 660
Db 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 1659

2Y 661 TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGACTGAACCACTATTGGAGCCTCT 720
Db 1660 TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGACTGAACCACTATTGGAGCCTCT 1719

2Y 721 CCTACTCAGACTGTTACTCTGGTGACAAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db 1720 CCTACTCAGACTGTTACTCTGGTGACAAACCTGTGGTTACTAAGGAAACTGCCATCTCC 1779

2Y 781 AAACCTAGAAATGCCATCTTCTTGTGTTGGAGGTACCTTACTCATAGATTTACTGCAACAG 840
Db 1780 AAACCTAGAAATGCCATCTTCTTGTGTTGGAGGTACCTTACTCATAGATTTACTGCAACAG 1839

2Y 841 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
Db 1840 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 1899

2Y 901 GTCCTACAGGATGCTACCCGTGAAGGAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAG 960
Db 901 GTCCTACAGGATGCTACCCGTGAAGGAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAG 960

Db 1900 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 1959
Qy 961 CTGATGAAACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGTCAACACAGATGTTTTATCAC 1020
Db 1960 CTGATGAAACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGTCAACACAGATGTTTTATCAC 2019
Qy 1021 AACCTGGATGAAACAGCCAAATAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTC 1080
Db 2020 AACCTGGATGAAACAGCCAAATAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTC 2079
Qy 1081 CTGTTACAAAGACGTTTGGATAACATGAACATTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140
Db 2080 CTGTTACAAAGACGTTTGGATAACATGAACATTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2139
Qy 1141 CTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 1200
Db 2140 CTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 2199
Qy 1201 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2200 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2259
Qy 1261 ATTGGAGGCGACTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCCTTCAAGAGG 1320
Db 2260 ATTGGAGGCGACTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCCTTCAAGAGG 2319
Qy 1321 GAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 1380
Db 2320 GAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 2379
Qy 1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCACAGAGCTGCCTCCT 1440
Db 2380 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCACAGAGCTGCCTCCT 2439
Qy 1441 GAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 2440 GAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 2499
Qy 1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 2500 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 2559
Qy 1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCT 1620
Db 2560 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCT 2619
Qy 1621 GAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
Db 2620 GAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 2679
Qy 1681 CACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAAATTGGCCTCTCTGAAAGAGAACCTGAGC 1740
Db 2680 CACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAAATTGGCCTCTCTGAAAGAGAACCTGAGC 2739
Qy 1741 CACGTCAATGACCTTGTGCGCCAGCTTACCACCTTTGGGCAFTCAGCTCTCACCCGTATAAC 1800
Db 2740 CACGTCAATGACCTTGTGCGCCAGCTTACCACCTTTGGGCAFTCAGCTCTCACCCGTATAAC 2799
Qy 1801 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGGTGGCCGTCGAGGAC 1860
Db 2800 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGGTGGCCGTCGAGGAC 2859
Qy 1861 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 2860 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTT 2919
Qy 1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCAAAACAAAGTGCCCTACTAT 1980
Db 2920 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCAAAACAAAGTGCCCTACTAT 2979
Qy 1981 ATCAACCACGAGACTCAACA 2001
Db 2980 ATCAACCACGAGACTCAACA 3000

ESULT 2
AD37256
D AAD37256 standard; DNA; 4966 BP.
X
C AAD37256;
X
X
T 21-AUG-2002 (first entry)
X
E Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
X
X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X
S Homo sapiens.
S Unidentified.
S Chimeric.
X
N WO200183695-A2.
D
X 08-NOV-2001.
X
X 27-APR-2001; 2001WO-US013677.
F
X 28-APR-2000; 2000US-0200777P.
R
X
A (XIAO/) XIAO X.
X
X Xiao X;
I
X WPI; 2002-049342/06.
R
X
T New dystrophin minigene for treating Duchenne or Becker muscular
I dystrophy comprises an N-terminal domain or modified N-terminal domain,
I rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
I gene.
X
S Example 1; Page 59-60; 71pp; English.
S
X
C The present invention relates to an isolated nucleotide sequence encoding
C a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is
C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C dystrophy (BMD) in a mammalian subject. The present sequence is AAV
C vector plasmid construct containing human dystrophin minigenes, a muscle
C creatine kinase (MCK) promoter and a small polyA signal sequence
X
Q Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 6; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
b
1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1816

Y 61 GTATTATCGTGGCTCTTTCTGCTGAGACACATTGCAAGCACAGGAGAGATTCTTAAT 120
b 1817 GTATTATCGTGGCTCTTTCTGCTGAGACACATTGCAAGCACAGGAGAGATTCTTAAT 1876

Y 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
b 1877 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1936

Y 181 GCCCATCAGGCGCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240

Db 1937 GCCCATCAGGCGCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGA 2056
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 420
Db 2117 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 2176
QY 421 ACAAGGAAAAATGGAGGAAGAGCCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 480
Db 2177 ACAAGGAAAAATGGAGGAAGAGCCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 2236
QY 481 CAACAACATAAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540
Db 2237 CAACAACATAAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 2296
QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 2297 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGACAGAAAGACCGC 660
Db 2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGACAGAAAGACCGC 2416
QY 661 TGGGTCTCTTTTACAAGACCAGCCCTGACCTAGTCTCTGGACTGACCACCTATTGGAGCCTCT 720
Db 2417 TGGGTCTCTTTTACAAGACCAGCCCTGACCTAGTCTCTGGACTGACCACCTATTGGAGCCTCT 2476
QY 721 CCTACTCAGACTGTTTACTCTGGTGACACACACCTGTGTTTACTAAGGAAACTGCCATCTCC 780
Db 2477 CCTACTCAGACTGTTTACTCTGGTGACACACACCTGTGTTTACTAAGGAAACTGCCATCTCC 2536
QY 781 AAACCTAGAAATGCCATCTCTTGTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 2537 AAACCTAGAAATGCCATCTCTTGTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2596
QY 841 TTCCCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAAACTGCCAAT 900
Db 2597 TTCCCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAAACTGCCAAT 2656
QY 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGCTCCAAGGGAGTAAAGAG 960
Db 2657 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGCTCCAAGGGAGTAAAGAG 2716
QY 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACAGAGATGTTTATCAC 1020
Db 2717 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACAGAGATGTTTATCAC 2776
QY 1021 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCGGATGTCAGTC 1080
Db 2777 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCGGATGTCAGTC 2836
QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140
Db 2837 CTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2896
QY 1141 CTCAAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAGTGGAAAGCGTCTGCACCTTTCT 1200
Db 2897 CTCAAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAGTGGAAAGCGTCTGCACCTTTCT 2956
QY 1201 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2957 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 3016
QY 1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAACAGAACCGATGTACATAGGGCCTTCAAGAGG 1320

3017 ATGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGG 3076
1321 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 1380
3077 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 3136
1381 ACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACTCTACCAAGGAGCCCGAGAGAGCTGCCCTCT 1440
3137 ACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACTCTACCAAGGAGCCCGAGAGAGCTGCCCTCT 3196
1441 GAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACT 1500
3197 GAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACT 3256
1501 GAGTGGGAAAAATTGAAACCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCCTT 1560
3257 GAGTGGGAAAAATTGAAACCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCCTT 3316
1561 GAAAGACTCCAGGAACCTTCAGAGGCCACCGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 1620
3317 GAAAGACTCCAGGAACCTTCAGAGGCCACCGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 3376
1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
3377 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 3436
1681 CACCTCGAGAAAGTCAAGGACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 1740
3437 CACCTCGAGAAAGTCAAGGACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 3496
1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
3497 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 3556
1801 CTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCCGTCGAGGAC 1860
3557 CTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCCGTCGAGGAC 3616
1861 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 1920
3617 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 3676
1921 TCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
3677 TCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3736
1981 ATCAACACGAGACTCAAAACA 2001
3737 ATCAACACGAGACTCAAAACA 3757
RESULT 3
D AAD37262 standard; DNA; 4990 BP.
C AAD37262;
C AAD37262;
T 21-AUG-2002 (first entry)
C Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
E Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
C adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
C Becker muscular dystrophy; ds.
C Homo sapiens.
S Cytomegalovirus.
S Unidentified.
S Chimeric.
C WO200183695-A2.
N 08-NOV-2001.
D

XX 27-APR-2001; 2001WO-US013677.
PF 28-APR-2000; 2000US-0200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
PI WPI; 2002-049342/06.
DR
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 67-68; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomagalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;
Query Match 100.0%; Score 2001; DB 6; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
DB 1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1840
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAAT 120
DB 1841 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAAT 1900
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGATTGACA 180
DB 1901 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGATTGACA 1960
QY 181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
DB 1961 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 2020
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
DB 2021 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2080
QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 2081 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2140
QY 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
DB 2141 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 2200
QY 421 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
DB 2201 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2260
QY 481 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAACCAAGTCAGGGTCAATTCTCTC 540
DB 2261 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAACCAAGTCAGGGTCAATTCTCTC 2320

541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
|||||
2321 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380
|||||
601 GAAACAACCTTAAGGTATTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660
|||||
2381 GAAACAACCTTAAGGTATTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2440
|||||
661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTAGCTGCTGGACTGACCACTATTGGAGCCTCT 720
|||||
2441 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTAGCTGCTGGACTGACCACTATTGGAGCCTCT 2500
|||||
721 CCTACTCAGACTGTTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 780
|||||
2501 CCTACTCAGACTGTTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 2560
|||||
781 AAACCTAGAAAATGCCATCTTCTCTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
|||||
2561 AAACCTAGAAAATGCCATCTTCTCTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2620
|||||
841 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTCCAAAGGAGTAAAGAG 900
|||||
2621 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTCCAAAGGAGTAAAGAG 2680
|||||
901 GTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAG 960
|||||
2681 GTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAG 2740
|||||
961 CTGATGAAACAATGGCAAGACCTTCCAAAGTGAATTTGAAGTCAACAGATGTTTATCAC 1020
|||||
2741 CTGATGAAACAATGGCAAGACCTTCCAAAGTGAATTTGAAGTCAACAGATGTTTATCAC 2800
|||||
1021 AACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGTTCGGATGATGCAGTC 1080
|||||
2801 AACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGTTCGGATGATGCAGTC 2860
|||||
1081 CTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 1140
|||||
2861 CTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 2920
|||||
1141 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 1200
|||||
2921 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 2980
|||||
1201 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 1260
|||||
2981 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 3040
|||||
1261 ATTGAGGCGACTTTTCCAGAGTTTCAAGAGCAACAGATGTACATAGGGCCTTCAAGAGG 1320
|||||
3041 ATTGAGGCGACTTTTCCAGAGTTTCAAGAGCAACAGATGTACATAGGGCCTTCAAGAGG 3100
|||||
1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTG 1380
|||||
3101 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTG 3160
|||||
1381 ACAGAGCAGCCTTTTGGAGGACTAGAGAAAACCTTACAGGAGCCCAAGAGAGCTGCCTCCT 1440
|||||
3161 ACAGAGCAGCCTTTTGGAGGACTAGAGAAAACCTTACAGGAGCCCAAGAGAGCTGCCTCCT 3220
|||||
1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTACGAAAAGCAGGCTGAGGAGGTCAATACT 1500
|||||
3221 GAGGAGAGAGCCAGAAATGTCACTCGGCTTACGAAAAGCAGGCTGAGGAGGTCAATACT 3280
|||||
1501 GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
|||||
3281 GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3340
|||||
1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 1620
|||||
3341 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 3400
|||||
1621 GAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGGGATCTCCTCATTGACTCTCTCCAAGAT 1680

Db 3401 GAGGTGATCAAGGGATCCTGGAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 3460
|||||
QY 1681 CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAAGAGAACGTGAGC 1740
|||||
Db 3461 CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAAGAGAACGTGAGC 3520
|||||
QY 1741 CACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 1800
|||||
Db 3521 CACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 3580
|||||
QY 1801 CTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGAC 1860
|||||
Db 3581 CTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGAC 3640
|||||
QY 1861 CGAGTCAGCAGCTGTCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 1920
|||||
Db 3641 CGAGTCAGCAGCTGTCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 3700
|||||
QY 1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT 1980
|||||
Db 3701 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT 3760
|||||
QY 1981 ATCAACCCAGAGACTCAAACA 2001
|||||
Db 3761 ATCAACCCAGAGACTCAAACA 3781

RESULT 4

AAD37237
ID AAD37237 standard; DNA; 3858 BP.

XX AAD37237;

AC AAD37237;

XX 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3849.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.

XX 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.

XX Example 1; Page 48-49; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is

useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)

Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match 85.4%; Score 1709; DB 6; Length 3858; Best Local Similarity 93.0%; Pred. No. 0; Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1059

61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1119

121 GATGTGGAAGTGGTGAAGACCAAGTTTTCATATCTCATGAGGGGTACATGATGGATTTGACA 180
1120 GATGTGGAAGTGGTGAAGACCAAGTTTTCATATCTCATGAGGGGTACATGATGGATTTGACA 1179

181 GCCCATCAGGGCCGGGTTGGTAAATATTCTACAAATTGGGAAGTAAAGTGAATGGAAACAGGA 240
1180 GCCCATCAGGGCCGGGTTGGTAAATATTCTACAAATTGGGAAGTAAAGTGAATGGAAACAGGA 1239

241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGCAGATGAATCTCCTAAATTCACAA 300
1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGCAGATGAATCTCCTAAATTCACAA 1299

301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATG 360
1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATG 1359

361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 420
1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 1419

421 ACAAGGAAAAATGGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTAAACCGCCAAAGTA 480
1420 ACAAGGAAAAATGGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTAAACCGCCAAAGTA 1479

481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTC 540
1480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTC 1539

541 ACTCACATGGTGGTGGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
1540 ACTCACATGGTGGTGGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599

601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1659

661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCTGGACTGACCACTATTGGAGCCTCT 720
1660 TGGGTTCTTTTACAAGAC----- 1677

721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 780
1678 ----- 1677

781 AAACATAGAAAATGCCATCTTCTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
1678 -----ACTCATAGATTACTGCAACAG 1698

841 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCCAAT 900
1699 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCCAAT 1758

901 GTCCTACAGGATGCTACCCCGTAAGGAAAGGCTCCTTAGAAGACTCCCAAGGGAGTAAAGAG 960

1759 GTCTACAGGATGCTACCCGTAAAGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAGAG 1818
961 CTGATGAAAAAATAATGGCAAGACCTCCAAGGTGAATTGAAGCTCACACAGATGTTTATCAC 1020
1819 CTGATGAAAAAATAATGGCAAGACCTCCAAGGTGAATTGAAGCTCACACAGATGTTTATCAC 1878

1021 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
1879 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1938

1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAATAAGTCT 1140
1939 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAATAAGTCT 1998

1141 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 1200
1999 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 2058

1201 CTGCAGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
2059 CTGCAGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2118

1261 ATTGGAGGCGACTTTCCAGAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGG 1320
2119 ATTGGAGGCGACTTTCCAGAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGG 2178

1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTCTG 1380
2179 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTCTG 2238

1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCCAGAGAGTGCCTCCT 1440
2239 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCCAGAGAGTGCCTCCT 2298

1441 GAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
2299 GAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 2358

1501 GAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTT 1560
2359 GAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTT 2418

1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
2419 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 2478

1621 GAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
2479 GAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 2538

1681 CACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGC 1740
2539 CACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGC 2598

1741 CAGGTCAATGACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
2599 CAGGTCAATGACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 2658

1801 CTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 1860
2659 CTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 2718

1861 CGAGTCAGGCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
2719 CGAGTCAGGCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 2778

1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 1980
2779 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 2838

1981 ATCAACCACGAGACTCAAAACA 2001

b 2839 ATCAACCAGAGACTCAAACA 2859

ESULT 5

AD37257

D AAD37257 standard; DNA; 4825 BP.

X X

C AAD37257;

X X

T 21-AUG-2002 (first entry)

X X

E Adeno-associated virus vector plasmid, AAV-MCK-delta3849.

X X

X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

W Becker muscular dystrophy; ds.

X X

S Homo sapiens.

S Unidentified.

S Chimeric.

X X

N WO200183695-A2.

X X

D 08-NOV-2001.

X X

F 27-APR-2001; 2001WO-US013677.

X X

R 28-APR-2000; 2000US-0200777P.

X X

A (XIAO/) XIAO X.

X X

I Xiao X;

X X

R WPI; 2002-049342/06.

X X

T New dystrophin minigene for treating Duchenne or Becker muscular

T dystrophy comprises an N-terminal domain or modified N-terminal domain,

T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

T gene.

X X

S Example 1; Page 61-62; 71pp; English.

X X

C The present invention relates to an isolated nucleotide sequence encoding

C a dystrophin minigene. The minigene comprises N-terminal or modified N-

C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

C domains and cysteine-rich domains of dystrophin or utrophin genes. The

C invention also relates to a recombinant adeno-associated virus (AAV)

C comprising dystrophin minigene operably linked to an expression control

C element. The dystrophin minigene in operable linkage with an expression

C control element, in a recombinant adeno-associated virus or retrovirus is

C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

C dystrophy (BMD) in a mammalian subject. The present sequence is AAV

C vector plasmid construct containing human dystrophin minigenes, a muscle

C creatine kinase (MCK) promoter and a small polyA signal sequence

X Q

Q Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match 85.4%; Score 1709; DB 6; Length 4825;

Best Local Similarity 93.0%; Pred. No. 0;

Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

Y 1 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60

b 1757 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1816

Y 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAAT 120

b 1817 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAAT 1876

Y 121 GATGTGGAAGTGGTGAAGACCAGTTTCTACTACTCATGAGGGGTACATGATGGATTGACA 180

b 1877 GATGTGGAAGTGGTGAAGACCAGTTTCTACTACTCATGAGGGGTACATGATGGATTGACA 1936

QY 181 GCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240

Db 1937 GCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 300

Db 1997 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 2056

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 360

Db 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 2116

QY 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAAAGAAAGA 420

Db 2117 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAAAGAAAGA 2176

QY 421 ACAAGGAAAAATGGAGGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 480

Db 2177 ACAAGGAAAAATGGAGGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 2236

QY 481 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 540

Db 2237 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 2296

QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600

Db 2297 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356

QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660

Db 2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2416

QY 661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCTGGACTGACCCTATTGAGCCTCT 720

Db 2417 TGGGTTCTTTTACAAGAC----- 2434

QY 721 CCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 780

Db 2435 ----- 2434

QY 781 AAACCTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840

Db 2435 -----ACTCATAGATTACTGCAACAG 2455

QY 841 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900

Db 2456 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2515

QY 901 GTCTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCCAAGGGAGTAAAAGAG 960

Db 2516 GTCTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCCAAGGGAGTAAAAGAG 2575

QY 961 CTGATGAAACAATGGCAAGACCTCCAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020

Db 2576 CTGATGAAACAATGGCAAGACCTCCAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2635

QY 1021 AACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080

Db 2636 AACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2695

QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 1140

Db 2696 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 2755

QY 1141 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 1200

Db 2756 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 2815

QY 1201 CTGAGGAACTTCTGTTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 1260

Db 2816 CTGAGGAACTTCTGTTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 2875

QY 1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320

Db	2876	ATTGGAGGGCACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG	29335
Qy	1321	GAATTGAAAACATAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAAATATTTCTG	1380
Db	2936	GAATTGAAAACATAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAAATATTTCTG	2995
Qy	1381	ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTTACAGGAGCCACAGAGAGCTGCCTCCT	1440
Db	2996	ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTTACAGGAGCCACAGAGAGCTGCCTCCT	3055
Qy	1441	GAGGAGAGAGCCACAGAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGGTCAATACT	1500
Db	3056	GAGGAGAGAGCCACAGAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGGTCAATACT	3115
Qy	1501	GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT	1560
Db	3116	GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT	3175
Qy	1561	GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT	1620
Db	3176	GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT	3235
Qy	1621	GAGGTGATCAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT	1680
Db	3236	GAGGTGATCAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT	3295
Qy	1681	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC	1740
Db	3296	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC	3355
Qy	1741	CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC	1800
Db	3356	CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC	3415
Qy	1801	CTCAGCACCTTGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC	1860
Db	3416	CTCAGCACCTTGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC	3475
Qy	1861	CGAGTCAGGCAGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTT	1920
Db	3476	CGAGTCAGGCAGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTT	3535
Qy	1921	TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT	1980
Db	3536	TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT	3595
Qy	1981	ATCAACCACGAGACTCAAACA	2001
Db	3596	ATCAACCACGAGACTCAAACA	3616

RESULT 6
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC
XX AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
DS Homo sapiens.
DS Cytomegalovirus.
DS Unidentified.
DS Chimeric.
XX
PN WO200183695-A2.
XX

08-NOV-2001.
 27-APR-2001; 2001WO-US013677.
 28-APR-2000; 2000US-0200777P.
 (XIAO/) XIAO X.
 Xiao X;
 WPI; 2002-049342/06.
 New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
 Example 1; Page 68-70; 71pp; English.
 The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a cytomegalovirus (CMV) promoter and a small polyA signal sequence
 Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;
 Query Match 85.4%; Score 1709; DB 6; Length 4848;
 Best Local Similarity 93.0%; Pred. No. 0;
 Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 17
 QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 60
 Db 1780 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 1839
 QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120
 Db 1840 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1899
 QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCTACTACTCATGAGGGGTACATGATGGATTGACA 180
 Db 1900 GATGTGGAAGTGGTGAAGACCAGTTTCTACTACTCATGAGGGGTACATGATGGATTGACA 1959
 QY 181 GCCCATCAGGCCGGGTGGTAAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
 Db 1960 GCCCATCAGGCCGGGTGGTAAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 2019
 QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACGAGATGAATCTCTAAATTCAGA 300
 Db 2020 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACGAGATGAATCTCTAAATTCAGA 2079
 QY 301 TGGGAATGCCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGATTTTAAATG 360
 Db 2080 TGGGAATGCCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGATTTTAAATG 2139
 QY 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
 Db 2140 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 2199
 QY 421 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 480
 Db 2200 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 2259
 QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC 540
 Db 2260 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC 2319

Y 541 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGAA 600
b 2320 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGAA 2379
Y 601 GAAACAACCTTAAGGTATGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACGC 660
b 2380 GAAACAACCTTAAGGTATGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACGC 2439
Y 661 TGGGTTCTTTTACAAGACAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 720
b 2440 TGGGTTCTTTTACAAGAC----- 2457
Y 721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 780
b 2458 ----- 2457
Y 781 AAAC TAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
b 2458 -----ACTCATAGATTACTGCAACAG 2478
Y 841 TTCCCCCTGGACCTGGAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
b 2479 TTCCCCCTGGACCTGGAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2538
Y 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTTCAAGGGAGTAAACAG 960
b 2539 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTTCAAGGGAGTAAACAG 2598
Y 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCAC 1020
b 2599 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCAC 2658
Y 1021 AACCTGGATGAAACAGGCAAAATAATCCTGAGATCCCTGGAAGTTCGGATGATGCACTC 1080
b 2659 AACCTGGATGAAACAGGCAAAATAATCCTGAGATCCCTGGAAGTTCGGATGATGCACTC 2718
Y 1081 CTGTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCT 1140
b 2719 CTGTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCT 2778
Y 1141 CTCAACATTAGGTCCTCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTCT 1200
b 2779 CTCAACATTAGGTCCTCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTCT 2838
Y 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACT 1260
b 2839 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACT 2898
Y 1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
b 2899 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 2958
Y 1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
b 2959 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3018
Y 1381 ACAGAGCAGCCTTTGGAAAGGACTAGAGAAATCTTACCAGGAGCCAGAGAGCTGCCCTCT 1440
b 3019 ACAGAGCAGCCTTTGGAAAGGACTAGAGAAATCTTACCAGGAGCCAGAGAGCTGCCCTCT 3078
Y 1441 GAGGAGAGAGCCCGAGATGTCTACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
b 3079 GAGGAGAGAGCCCGAGATGTCTACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3138
Y 1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTCCGCTGAGTGGCAGAGAAAAATAGATGAGACCCCTT 1560
b 3139 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3198
Y 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 1620
b 3199 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 3258

QY 1621 GAGTGATCAAGGGATCTTGCGCAGCCCGTGGCGCATCTCTCATTTGACTCTCTCCAAGAT 1680
Db 3259 GAGTGATCAAGGGATCTTGCGCAGCCCGTGGCGCATCTCTCATTTGACTCTCTCCAAGAT 3318
QY 1681 CACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGC 1740
Db 3319 CACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGC 3378
QY 1741 CACGTCAATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAC 1800
Db 3379 CACGTCAATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAC 3438
QY 1801 CTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGTTCGAGGAC 1860
Db 3439 CTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGTTCGAGGAC 3498
QY 1861 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 1920
Db 3499 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 3558
QY 1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
Db 3559 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3618
QY 1981 ATCAACACGAGACTCAAAACA 2001
Db 3619 ATCAACACGAGACTCAAAACA 3639
RESULT 7
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX
AC AAD37264;
XX 21-AUG-2002 (first entry)
XX Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a small polyA signal sequence

SQ	Sequence	5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;
	Query Match	85.4%; Score 1709; DB 6; Length 5060;
	Best Local Similarity	93.0%; Pred. No. 0;
	Matches 1860; Conservative	0; Mismatches 0; Indels 141; Gaps 1;
QY	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTAGAAGAA 60
Db	1992	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTAGAAGAA 2051
QY	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 120
Db	2052	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 2111
QY	121	GATGTGGAAGTGTGAAAGACCAGTTTCATACTACTAGGGGTACATGATGGATTGACA 180
Db	2112	GATGTGGAAGTGTGAAAGACCAGTTTCATACTACTAGGGGTACATGATGGATTGACA 2171
QY	181	GCCCATCAGGGCCGGTGGTAATATCTACAATTGGGAAGTAAGCTGATTGAAACAGGA 240
Db	2172	GCCCATCAGGGCCGGTGGTAATATCTACAATTGGGAAGTAAGCTGATTGAAACAGGA 2231
QY	241	AAATTATCAGAAAGATGAAGAACTGAACTCAAGACAGATGAATCTCCTAAATCAAGA 300
Db	2232	AAATTATCAGAAAGATGAAGAACTGAACTCAAGACAGATGAATCTCCTAAATCAAGA 2291
QY	301	TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 360
Db	2292	TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 2351
QY	361	GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 420
Db	2352	GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 2411
QY	421	ACAAGGAAATCGAGGAAGAGCCCTCTTGGACCTTGATCTTGAAGACCTAAACGCCAAGTA 480
Db	2412	ACAAGGAAATCGAGGAAGAGCCCTCTTGGACCTTGATCTTGAAGACCTAAACGCCAAGTA 2471
QY	481	CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC 540
Db	2472	CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC 2531
QY	541	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db	2532	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2591
QY	601	GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660
Db	2592	GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2651
QY	661	TGGGTTCCTTTACAAGACCAGCCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCCTCT 720
Db	2652	TGGGTTCCTTTACAAGAC----- 2669
QY	721	CCTACTCAGACTGTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db	2670	----- 2669
QY	781	AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db	2670	-----ACTCATAGATTACTGCAACAG 2690
QY	841	TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAAACAACACTGCCAAT 900

Db	2691	TTCCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAAACAACCTGCCAAT 2750
QY	901	GTCTACAGGATGTACCCGTAAGGAAAGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 960
Db	2751	GTCTACAGGATGTACCCGTAAGGAAAGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 2810
QY	961	CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
Db	2811	CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2870
QY	1021	AACCTGGATGAAAACAGCCAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTC 1080
Db	2871	AACCTGGATGAAAACAGCCAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTC 2930
QY	1081	CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGA AAAAGTCT 1140
Db	2931	CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGA AAAAGTCT 2990
QY	1141	CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 1200
Db	2991	CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 3050
QY	1201	CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db	3051	CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 3110
QY	1261	ATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
Db	3111	ATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3170
QY	1321	GAAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db	3171	GAAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3230
QY	1381	ACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAAACCTTACCAGGAGGCCAGAGAGCTGCCTCCT 1440
Db	3231	ACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAAACCTTACCAGGAGGCCAGAGAGCTGCCTCCT 3290
QY	1441	GAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db	3291	GAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3350
QY	1501	GAGTGGGAAAAAATTGAAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db	3351	GAGTGGGAAAAAATTGAAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3410
QY	1561	GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCT 1620
Db	3411	GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCT 3470
QY	1621	GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
Db	3471	GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 3530
QY	1681	CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGC 1740
Db	3531	CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGC 3590
QY	1741	CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 1800
Db	3591	CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 3650
QY	1801	CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 1860
Db	3651	CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 3710
QY	1861	CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db	3711	CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3770
QY	1921	TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTAT 1980

b 3771 TCACGCTGTGCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3830
y 1981 ATCAACCACGAGACTCAAACA 2001
b 3831 ATCAACCACGAGACTCAAACA 3851
RESULT 8
AD37230
D AAD37230 standard; DNA; 4182 BP.
X
C AAD37230;
X
T 21-AUG-2002 (first entry)
X
E Human dystrophin minigene delta4173.
X
W Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
W
S Homo sapiens.
X
X WO200183695-A2.
N
X
X 08-NOV-2001.
X
F 27-APR-2001; 2001WO-US013677.
X
R 28-APR-2000; 2000US-0200777P.
X
A (XIAO/) XIAO X.
X
I Xiao X;
X
R WPI; 2002-049342/06.
X
X
T New dystrophin minigene for treating Duchenne or Becker muscular
T dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.
X
S Example 1; Page 43-44; 71pp; English.
X
C The present invention relates to an isolated nucleotide sequence encoding
C a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is
C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C dystrophy (BMD) in a mammalian subject. The present sequence is human
C dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
C hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
C hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
X
Q Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
Query Match 83.1%; Score 1662.8; DB 6; Length 4182;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
y 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
b 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1059
y 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAAT 120
b 1060 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAAT 1119
y 121 GATGTGGAAGTGGTGAAGAACCAAGTTCATACTCATGAGGGGACATGATGGAATTGACA 180

Db 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1179
QY 181 GCCCATCAGGGCCGGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1180 GCCCATCAGGGCCGGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTTCAAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGA 420
Db 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGA 1419
QY 421 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAGTA 480
Db 1420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAGTA 1479
QY 481 CAACAAACATAAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 540
Db 1480 CAACAAACATAAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 1539
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGC 660
Db 1600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGC 1659
QY 661 TGGGTTCTTTTACAAGACCAGCCTGAC-----CTAGCTCTCTGGACTGACCACT 708
Db 1660 TGGGTTCTTTTACAAGACATCTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTT 1719
QY 709 ATTGAGCCCTCTCCTACTCAGACTGTTACTCTGTGTGACACA-----ACCTGTGGTT 759
Db 1720 TTTAGTGCATGGCTTTCAGAAAAAAGAGATGCGAGTGAACAAAGATTCAACAACACTGGCTTT 1779
QY 760 ACTAAGGAAAACTGCCATCTC----- 779
Db 1780 AAAGATCAAAATGAAATGTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAAGCGGATCTA 1839
QY 780 -----CAAACCTAGAAATGCCATCTTCC 801
Db 1840 GAAAAAGAAAAAGCAATCCATGGGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACA 1899
QY 802 TTGATGTTGGAG----- 813
Db 1900 CTGAAGAATAAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT 1959
QY 814 -----GTACCTACTCATAGATTACTGCAA 837
Db 1960 TGGGATAATTTAGTCCAAAAAACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019
QY 838 CAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAAACAACTGCC 897
Db 2020 CAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAAACAACTGCC 2079
QY 898 AATGTCCTACAGGATGCTACCCGTAAGGAAAGCTCCTAGAAAGCTCCAAAGGGAGTAAAA 957
Db 2080 AATGTCCTACAGGATGCTACCCGTAAGGAAAGCTCCTAGAAAGCTCCAAAGGGAGTAAAA 2139
QY 958 GAGCTGATGAACAATGCAAGACCTCCCAAGTGAATTTGAAGCTCACACAGATGTTTAT 1017
Db 2140 GAGCTGATGAACAATGCAAGACCTCCCAAGTGAATTTGAAGCTCACACAGATGTTTAT 2199
QY 1018 CACAACCTGGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCA 1077

Db 2200 CACAACTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCA 2259
2Y 1078 GTCCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAG 1137
Db 2260 GTCCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAG 2319
2Y 1138 TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTT 1197
Db 2320 TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTT 2379
2Y 1198 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCA 1257
Db 2380 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCA 2439
2Y 1258 CCTATTGGAGCGCACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGSCCTTCAAG 1317
Db 2440 CCTATTGGAGCGCACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGSCCTTCAAG 2499
2Y 1318 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 1377
Db 2500 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 2559
2Y 1378 CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGCAGAGAGCTGCCT 1437
Db 2560 CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGCAGAGAGCTGCCT 2619
2Y 1438 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTTCAAT 1497
Db 2620 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTTCAAT 2679
2Y 1498 ACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db 2680 ACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
2Y 1558 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAA 1617
Db 2740 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAA 2799
2Y 1618 GCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 1677
Db 2800 GCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 2859
2Y 1678 GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTG 1737
Db 2860 GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTG 2919
2Y 1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 1797
Db 2920 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 2979
2Y 1798 AACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAG 1857
Db 2980 AACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAG 3039
2Y 1858 GACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTT 1917
Db 3040 GACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTT 3099
2Y 1918 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTAC 1977
Db 3100 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTAC 3159
2Y 1978 TATATCAACACGAGACTCAAAACA 2001
Db 3160 TATATCAACACGAGACTCAAAACA 3183

RESULT 9
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX

DT 21-AUG-2002 (first entry)
XX Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX WO200183695-A2.
PN 08-NOV-2001.
XX 27-APR-2001; 2001WO-US013677.
PF 28-APR-2000; 2000US-0200777P.
XX (XIAO/) XIAO X.
PA Xiao X;
XX WPI; 2002-049342/06.
DR New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 57-59; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;
Query Match 83.1%; Score 1662.8; DB 6; Length 5149;
Best Local Similarity 88.3%; Pred.No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCAGTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCAGTTCATCTCATGAGGGGTACATGATGGATTGACA 1936
QY 181 GCCCATCAGGCCCGGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
Db 1937 GCCCATCAGGCCCGGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2056
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360

b 2057 TGGGAATGCCTCAGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTTAAATG 2116
y 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 420
b 2117 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 2176
y 421 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAAGCCCAAGTA 480
b 2177 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAAGCCCAAGTA 2236
y 481 CAACAACATAAGGTGTTCAAGAAGATCTAGAACAAGAACAAGTCAAGGTCAATTTCTCTC 540
b 2237 CAACAACATAAGGTGTTCAAGAAGATCTAGAACAAGAACAAGTCAAGGTCAATTTCTCTC 2296
y 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
b 2297 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356
y 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 660
b 2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 2416
y 661 TGGGTTCTTTTACAGACCAAGCCTGAC-----CTAGCTCCTGGACTGACCACT 708
b 2417 TGGGTTCTTTTACAGACATCTTCTCAATGGCAACGCTTCTTACTGAAGAACAGTGCCTT 2476
y 709 ATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACA-----ACCTGTGTT 759
b 2477 TTTAGTGATGGCTTTCAGAAAAAGAAAGATGCAGTGAACAAGATTACACAACCTGGCTTT 2536
y 760 ACTAAGGAAACTGCCATCTC-----779
b 2537 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTA 2596
y 780 -----CAAAGTAAATGCCATCTTCC 801
b 2597 GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 2656
y 802 TTGATGTTGGAG-----813
b 2657 CTGAAGAATAAGTCACTGACCCAGAACGGAAGCATGGCTGGATAAACTTTGCCCGGTGT 2716
y 814 -----GTACCTACTCATAGATTACTGCAA 837
b 2717 TGGGATAATTTAGTCCAAAAAACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
y 838 CAGTTCCCTCCCTGGACTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCC 897
b 2777 CAGTTCCCTCCCTGGACTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCC 2836
y 898 AATGTCCTACAGGATGTAACCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAA 957
b 2837 AATGTCCTACAGGATGTAACCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAA 2896
y 958 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTAT 1017
b 2897 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTAT 2956
y 1018 CACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 1077
b 2957 CACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 3016
y 1078 GTCTGTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAG 1137
b 3017 GTCTGTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAG 3076
y 1138 TCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCAGTGAAGCGGTCTGCACCTT 1197
b 3077 TCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCAGTGAAGCGGTCTGCACCTT 3136
y 1198 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCA 1257

Db 3137 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCA 3196
Qy 1258 CCTATTGGAGCGGACTTTCCAGCAGTTTCAGAACGAGAACGATGTACATAGGCGCTTCAAG 1317
Db 3197 CCTATTGGAGCGGACTTTCCAGCAGTTTCAGAACGAGAACGATGTACATAGGCGCTTCAAG 3256
Qy 1318 AGGGAATTGAAAACCTAAAGAACCTGTAATCATGACTACTCTTTGAGACTGTACGAATATTT 1377
Db 3257 AGGGAATTGAAAACCTAAAGAACCTGTAATCATGACTACTCTTTGAGACTGTACGAATATTT 3316
Qy 1378 CTGACAGAGCAGCCCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCT 1437
Db 3317 CTGACAGAGCAGCCCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCT 3376
Qy 1438 CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTTTACGAAAGCAGGCTGAGGAGGTCAAT 1497
Db 3377 CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTTTACGAAAGCAGGCTGAGGAGGTCAAT 3436
Qy 1498 ACTGAGTGGGAAAAAATTGAACTTCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db 3437 ACTGAGTGGGAAAAAATTGAACTTCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
Qy 1558 CTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA CCTCAAGCTGCGCCAA 1617
Db 3497 CTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA CCTCAAGCTGCGCCAA 3556
Qy 1618 GCTGAGTGATCAAGGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 1677
Db 3557 GCTGAGTGATCAAGGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 3616
Qy 1678 GATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTG 1737
Db 3617 GATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTG 3676
Qy 1738 AGCCAGCTCAATGACCTTGCTCGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCCGTAT 1797
Db 3677 AGCCAGCTCAATGACCTTGCTCGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCCGTAT 3736
Qy 1798 AACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAG 1857
Db 3737 AACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAG 3796
Qy 1858 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTT 1917
Db 3797 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTT 3856
Qy 1918 CTTTCCACGTCGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAGTGCCTTAC 1977
Db 3857 CTTTCCACGTCGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAGTGCCTTAC 3916
Qy 1978 TATATCAACCACGAGACTCAAAACA 2001
Db 3917 TATATCAACCACGAGACTCAAAACA 3940

RESULT 10
ABK81999
ID ABK81999 standard; DNA; 5462 BP.
XX
AC ABK81999;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
XX
PN Synthetic.
XX
WC200229056-A2. 80

PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Disclosure; Fig 14; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;
Query Match 64.1%; Score 1283; DB 6; Length 5462;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
691 GCTCCTGGACTGACCACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGTGACACAA 750
1547 GCTCCTGGACTGACCACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGTGACACAA 1606
751 CCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 810
1607 CCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 1666
811 GAGGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAGTTCTTGCC 870
1667 GAG-----CATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAGTTCTTGCC 1717
871 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGGAAAGG 930
1718 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGGAAAGG 1777
931 CTCCTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGGT 990
1778 CTCCTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGGT 1837
991 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATCCTG 1050
1838 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATCCTG 1897
1051 AGATCCCTGGAAGGTTCCGATGATGCAACCTGTTTACAAAGACGTTTGGATAACATGAAC 1110
1898 AGATCCCTGGAAGGTTCCGATGATGCAACCTGTTTACAAAGACGTTTGGATAACATGAAC 1957
1111 TTCAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGT 1170
1958 TTCAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGT 2017
1171 TCTGACCAGTGAAGGTTCTGCACCTTTCTCTGCAGGAACCTTCTGTTGGTGGCTACAGCTG 1230
2018 TCTGACCAGTGAAGGTTCTGCACCTTTCTCTGCAGGAACCTTCTGTTGGTGGCTACAGCTG 2077
1231 AAAGATGATGAATTAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAG 1290

Db 2078 AAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCCAGAAG 2137
QY 1291 CAGAACGATGTACATAGGGCCCTTCAAGAGAGGAATTTGAAAACTAAAGAACCTGTATCATG 1350
Db 2138 CAGAACGATGTACATAGGGCCCTTCAAGAGAGGAATTTGAAAACTAAAGAACCTGTATCATG 2197
QY 1351 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 1410
Db 2198 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 2257
QY 1411 CTCTACCAGGAGCCAGAGAGCTGCCCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTT 1470
Db 2258 CTCTACCAGGAGCCAGAGAGCTGCCCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTT 2317
QY 1471 CTACGAAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCCTCCGCT 1530
Db 2318 CTACGAAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCCTCCGCT 2377
QY 1531 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG 1590
Db 2378 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG 2437
QY 1591 GATGAGCTGGACCTCAAGCTGCGCCAAAGTGAGGTGATCAAGGGATCTTGGCAGCCCGTG 1650
Db 2438 GATGAGCTGGACCTCAAGCTGCGCCAAAGTGAGGTGATCAAGGGATCTTGGCAGCCCGTG 2497
QY 1651 GCGGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTTCGAGGA 1710
Db 2498 GCGGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTTCGAGGA 2557
QY 1711 GAAATTGCGCCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACC 1770
Db 2558 GAAATTGCGCCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACC 2617
QY 1771 ACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGA 1830
Db 2618 ACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGA 2677
QY 1831 TGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGG 1890
Db 2678 TGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGG 2737
QY 1891 GACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTCCTCGGGAGAGA 1950
Db 2738 GACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTCCTCGGGAGAGA 2797
QY 1951 GCCATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
Db 2798 GCCATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2848
RESULT 11
AAD06794
ID AAD06794 standard; DNA; 5952 BP.
XX
AC AAD06794;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human dystrophin gene (Becker form).
XX
KW Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_recomb 2847..2848
FT /*tag= a
FT /label= S4_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting fragment is
FT used to produce plasmid pSD4"

misc_recomb 2952. .2953
/*tag= b
/label= S3_junction_site
/note= "Dystrophin gene is split at this site and
attached to an intein sequence. The resulting fragment is
used to produce plasmid pSD3"
3198. .3199
/*tag= c
/label= S2_junction_site
/note= "Dystrophin gene is split at this site and
attached to an intein sequence. The resulting fragment is
used to produce plasmid pSD2"
3300. .3301
/*tag= d
/label= S1_junction_site
/note= "Dystrophin gene is split at this site and
attached to an intein sequence. The resulting fragment is
used to produce plasmid pSD1"

WO200129243-A1.
26-APR-2001.
13-OCT-2000; 2000WO-CA001216.
15-OCT-1999; 99US-0159868P.
(UYDA-) UNIV DALHOUSIE.
(UYPI-) UNIV PITTSBURGH.
Paul XL, Xiao X;
WPI; 2001-367297/38.
Use of spontaneous or automatic protein splicing to join two or more
peptides at junction site involves expressing extein peptides having co-
reacting portions of split intein attached to them, so that peptides
splice.

Example 1; Fig 2; 81pp; English.
The invention relates to a method directed to the use of spontaneous or
automatic protein trans-splicing to join two or more peptides at junction
site. This method involves expressing extein peptides having co-reacting
portions of split intein attached to them, so that peptides will splice
automatically under suitable conditions. The invention also provides
methods for circumventing virion packaging size limitations in
recombinant virus particle, by splitting a coding region for a protein to
be delivered into two or more extein genes, which are packaged in
separate virus particles and are co-delivered in a target cell for the
expression and for subsequent trans-splicing to form the complete
protein. In particular, the method is used for trans-splicing human
dystrophin and in gene therapies of recombinant adeno-associated virus
(AAV) particles that encode trans-spliced dystrophin, for treating
diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
dystrophy (BMD). The present sequence is human dystrophin gene (Becker
form)

Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;
Query Match 59.1%; Score 1182.6; DB 5; Length 5952;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTCTTGCCTG 872
|||||
2946 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTCTTGCCTG 3005
|||||

873 GCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGCT 932
|||||

3006 GCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGCT 3065
|||||

933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGATAAACAATGGCAAGACTCCAAGGTGA 992

Db 3066 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGATAAACAATGGCAAGACTCCAAGGTGA 3125
|||
Qy 993 AATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAATCCTGAG 1052
|||||
Db 3126 AATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAATCCTGAG 3185
|||||
Qy 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTTACAAAGACGTTTGGATAACATGAACCTT 1112
|||||
Db 3186 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTTACAAAGACGTTTGGATAACATGAACCTT 3245
|||||
Qy 1113 CAAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTC 1172
|||||
Db 3246 CAAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTC 3305
|||||
Qy 1173 TGACCAGTGAAGCGTCTGCACCTTTCTGTCAGGAACCTTGGTGTGGCTACAGCTGAA 1232
|||||
Db 3306 TGACCAGTGAAGCGTCTGCACCTTTCTGTCAGGAACCTTGGTGTGGCTACAGCTGAA 3365
|||||
Qy 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 1292
|||||
Db 3366 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 3425
|||||
Qy 1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAATAAAGAACCTGTATCATGAG 1352
|||||
Db 3426 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAATAAAGAACCTGTATCATGAG 3485
|||||
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 1412
|||||
Db 3486 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 3545
|||||
Qy 1413 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGCTTCT 1472
|||||
Db 3546 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGCTTCT 3605
|||||
Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCGCTGA 1532
|||||
Db 3606 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCGCTGA 3665
|||||
Qy 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
|||||
Db 3666 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 3725
|||||
Qy 1593 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGGATCCTGGCAGCCCGTGGG 1652
|||||
Db 3726 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGGATCCTGGCAGCCCGTGGG 3785
|||||
Qy 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGA 1712
|||||
Db 3786 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGA 3845
|||||
Qy 1713 AATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
|||||
Db 3846 AATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905
|||||
Qy 1773 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATG 1832
|||||
Db 3906 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATG 3965
|||||
Qy 1833 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
|||||
Db 3966 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 4025
|||||
Qy 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGC 1952
|||||
Db 4026 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGC 4085
|||||
Qy 1953 CATCTCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
|||||
Db 4086 CATCTCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4134
|||||

ABK82000
ID ABK82000 standard; DNA; 8689 BP.
XX
AC ABK82000;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200229056-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US031126.
XX
XX 06-OCT-2000; 2000US-0238848P.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains, or
XX a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
XX Disclosure; Fig 15; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
XX comprises n spectrin-like repeats, and contains no more than n spectrin-
XX like repeats, where n is an even number between 4-24, or a nucleic acid
XX encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
XX polynucleotide encoding it is useful as a medicament, for preparing a
XX drug for therapeutic application and in the preparation of a composition
XX for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
XX (DMD). This sequence represents a mini-dystrophin sequence of the
XX invention
SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;
Query Match 59.1%; Score 1182.6; DB 6; Length 8689;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAGTTTCTTGCTG 872
Db 2992 GGAAGAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAGTTTCTTGCTG 3051
QY 873 GCTTACAGAAGCTGAACAACCTGCTTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db 3052 GCTTACAGAAGCTGAACAACCTGCTTACAGGATGCTACCCGTAAGGAAAGGCT 3111
QY 933 CCTAGAAGACTCCAAGGAGTAAAGAGCTGATGAACAATGSCAAGACCTCCAAGTGA 992
Db 3112 CCTAGAAGACTCCAAGGAGTAAAGAGCTGATGAACAATGSCAAGACCTCCAAGTGA 3171
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCCTGAG 1052
Db 3172 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCCTGAG 3231
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAATT 1112
Db 3232 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAATT 3291
QY 1113 CAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172

Db 3292 CAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 3351
QY 1173 TGACCCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 3352 TGACCCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3411
QY 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGGCAGCTTTCCACGAGTTTCAAGCA 1292
Db 3412 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGGCAGCTTTCCACGAGTTTCAAGCA 3471
QY 1293 GAACCATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACCTTAAGAACCTTGTAAATCATGAG 1352
Db 3472 GAACCATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACCTTGTAAATCATGAG 3531
QY 1353 TACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGAGGACTAGAGAAACT 1412
Db 3532 TACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGAGGACTAGAGAAACT 3591
QY 1413 CTACCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 3592 CTACCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 3651
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA 1532
Db 3652 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA 3711
QY 1533 CTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGA 1592
Db 3712 CTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGA 3771
QY 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGAATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 3772 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGAATCAAGGGATCCTGGCAGCCCGTGGG 3831
QY 1653 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 1712
Db 3832 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 3891
QY 1713 AATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCAC 1772
Db 3892 AATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCAC 3951
QY 1773 TTTGGGCATTTGAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAAACACCATG 1832
Db 3952 TTTGGGCATTTGAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAAACACCATG 4011
QY 1833 GAAGTTTCTGCAGGTGGCGCTGAGGACCGAGTCCAGGAGCTGCATGAAGCCACAGGGA 1892
Db 4012 GAAGTTTCTGCAGGTGGCGCTGAGGACCGAGTCCAGGAGCTGCATGAAGCCACAGGGA 4071
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGC 1952
Db 4072 CTTTGGTCCAGCATCTCAGCACTTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGC 4131
QY 1953 CATCTGCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAACA 2001
Db 4132 CATCTGCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAACA 4180
RESULT 13
AAD37229
ID AAD37229 standard; DNA; 11058 BP.
XX
AC AAD37229;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin protein coding sequence.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Example 2; Fig 23; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a human dystrophin polynucleotide
CC sequence used in the creation of the mini-dystrophin peptides of the
CC invention
XX
SQ Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 U; 0 Other;
Query Match 59.1%; Score 1182.6; DB 6; Length 11241;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCTG 872
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCTG 8319
QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTGAAGAAAGGCT 932
Db 8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTGAAGAAAGGCT 8379
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
Db 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAGCCAAAAAATCCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAGCCAAAAAATCCTGAG 8499
QY 1053 ATCCCTGGAAGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAATT 1112
Db 8500 ATCCCTGGAAGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAATT 8559
QY 1113 CAAGTGGAGTGAACITCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
Db 8560 CAAGTGGAGTGAACITCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 8619
QY 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAAGCA 1292
Db 8680 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAAGCA 8739
QY 1293 GAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
Db 8740 GAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATCATGAG 8799
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8859
QY 1413 CTACAGAGAGCCAGAGAGTGCCTCCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 8860 CTACAGAGAGCCAGAGAGTGCCTCCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 8919

QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCCTCGCTGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCCTCGCTGA 8979
QY 1533 CTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGAACTCCAGGAACTTCAAGAGGCCACGGA 1592
Db 8980 CTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGAACTCCAGGAACTTCAAGAGGCCACGGA 9039
QY 1593 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCTGGG 1652
Db 9040 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCTGGG 9099
QY 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 1712
Db 9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 9159
QY 1713 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGCGCCAGCTTACCAC 1772
Db 9160 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGCGCCAGCTTACCAC 9219
QY 1773 TTTGGGCATTTCAGCTCTCACCCTGATATAACCTCAGCACTCTGGAAGACCTGAAACACCATG 1832
Db 9220 TTTGGGCATTTCAGCTCTCACCCTGATATAACCTCAGCACTCTGGAAGACCTGAAACACCATG 9279
QY 1833 GAAGCTTCTGCAGGTGCGCTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 1892
Db 9280 GAAGCTTCTGCAGGTGCGCTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 9339
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCTGGGAGAGAGC 1952
Db 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCTGGGAGAGAGC 9399
QY 1953 CATCTGCCCCAAACAAAGTGCCTTACTATATATCAACCCAGAGACTCAAACA 2001
Db 9400 CATCTGCCCCAAACAAAGTGCCTTACTATATATCAACCCAGAGACTCAAACA 9448

RESULT 15
ABK82002
ID ABK82002 standard; DNA; 11443 BP.
XX
AC ABK82002;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Disclosure; Fig 17; 145pp; English.

10 The invention describes a composition comprising a mini-dystrophin
 11 peptide comprising a spectrin-like repeat domain, where the domain
 12 comprises n spectrin-like repeats, and contains no more than n spectrin-
 13 like repeats, where n is an even number between 4-24, or a nucleic acid
 14 encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
 15 polynucleotide encoding it is useful as a medicament, for preparing a
 16 drug for therapeutic application and in the preparation of a composition
 17 for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
 18 (DMD). This sequence represents a mini-dystrophin sequence of the
 19 invention
 20
 21 Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;
 22
 23 Query Match 59.1%; Score 1182.6; DB 6; Length 11443;
 24 Best local Similarity 99.7%; Pred. No. 0;
 25 Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 26
 27 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
 28
 29 5746 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 5805
 30
 31 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
 32
 33 5806 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 5865
 34
 35 933 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 992
 36
 37 5866 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 5925
 38
 39 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAACAAACAGCCAAATAATCCTGAG 1052
 40
 41 5926 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAACAAACAGCCAAATAATCCTGAG 5985
 42
 43 1053 ATCCCTGGAAGTTCCGATGATGCGAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
 44
 45 5986 ATCCCTGGAAGTTCCGATGATGCGAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCTT 6045
 46
 47 1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTC 1172
 48
 49 6046 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTC 6105
 50
 51 1173 TGACCACTGGAAGCGTCTGCACCTTTCTGACAGGAACTTCTGGTGTGGCTACAGCTGAA 1232
 52
 53 6106 TGACCACTGGAAGCGTCTGCACCTTTCTGACAGGAACTTCTGGTGTGGCTACAGCTGAA 6165
 54
 55 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 1292
 56
 57 6166 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 6225
 58
 59 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
 60
 61 6226 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAG 6285
 62
 63 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
 64
 65 6286 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 6345
 66
 67 1413 CTACCCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCT 1472
 68
 69 6346 CTACCCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCT 6405
 70
 71 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 1532
 72
 73 6406 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 6465
 74
 75 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
 76
 77 6466 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 6525
 78
 79 1593 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
 80
 81 6526 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 6585

Qy 1653 CGATCTCCTCATGTACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
 Db
 6586 CGATCTCCTCATGTACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 6645
 Qy 1713 AATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
 Db 6646 AATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 6705
 Qy 1773 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAAACACCCAGATG 1832
 Db 6706 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAAACACCCAGATG 6765
 Qy 1833 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
 Db 6766 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 6825
 Qy 1893 CTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
 Db 6826 CTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 6885
 Qy 1953 CATCTCGCCAAACAAAAGTGCCCTACTATATCAACCCACGAGACTCAAAACA 2001
 Db 6886 CATCTCGCCAAACAAAAGTGCCCTACTATATCAACCCACGAGACTCAAAACA 6934

Search completed: April 4, 2004, 18:19:39
 Job time : 529.337 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 nucleic - nucleic search, using sw model
in on: April 4, 2004, 12:04:07 ; Search time 5216.76 Seconds
(without alignments)
16625.161 Million cell updates/sec
title: US-09-845-416-6_COPY_1000_3000
perfect score: 2001
sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaca 2001
scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
sarched: 3470272 seqs, 21671516995 residues
otal number of hits satisfying chosen parameters: 6940544
linum DB seq length: 0
ximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1283	64.1	5462	6	AX538621	AX538621 Sequence
2	1182.6	59.1	5952	6	AX304538	AX304538 Sequence
3	1182.6	59.1	5952	6	AX114289	AX114289 Sequence
4	1182.6	59.1	8689	6	AX538622	AX538622 Sequence
5	1182.6	59.1	11443	6	AX538624	AX538624 Sequence
6	1182.6	59.1	12057	6	AX538627	AX538627 Sequence
7	1182.6	59.1	12446	9	HSDMDR	X14298 Human mRNA
8	1182.6	59.1	13957	6	AX409637	AX409637 Sequence
9	1182.6	59.1	13957	6	AX538581	AX538581 Sequence
10	1182.6	59.1	13957	9	HUMDYS	M18533 Homo sapien
11	1182.6	59.1	13977	6	AR220819	AR220819 Sequence
12	1180	59.0	5339	6	AX538620	AX538620 Sequence
13	1030.6	51.5	13887	4	AF070485	AF070485 Canis fam
14	1004	50.2	13815	6	AX306153	AX306153 Sequence
15	1004	50.2	13815	6	AX538582	AX538582 Sequence
16	1004	50.2	13815	10	MUSDYSA	M68859 Mouse dystro
17	1004	50.2	19307	6	AR093392	AR093392 Sequence
18	1004	50.2	19307	6	AR142592	AR142592 Sequence
19	765	38.2	13575	5	GGDYS	X13369 Chicken mRN
20	727	36.3	5417	6	AX538619	AX538619 Sequence
21	678.4	33.9	4402	6	E30220	E30220 Shortened d
22	618.8	30.9	4075	6	E30221	E30221 Shortened d
23	614.2	30.7	4402	6	E30219	E30219 Shortened d
24	547.4	27.4	3275	10	MUSDYS	M18025 Mouse dystro
25	459.4	23.0	530	9	HSDMDFL	X06179 Human fetal
26	422.6	21.1	3521	5	AF339031	AF339031 Danio rer
27	414.6	20.7	1481	9	BC009242	BC009242 Homo sapi
28	408.2	20.4	11096	6	AX538584	AX538584 Sequence
29	408.2	20.4	11096	10	MMY12229	Y12229 M.musculus
30	408	20.4	10705	10	RNAJ2967	AJ002967 Rattus no
31	397	19.8	6045	6	A63605	A63605 Sequence 7
32	397	19.8	6045	6	AR281528	AR281528 Sequence
33	397	19.8	6059	6	AX107972	AX107972 Sequence
34	397	19.8	10302	6	AX538583	AX538583 Sequence
35	397	19.8	10302	9	HSMUPS	X69086 H.sapiens m
36	397	19.8	10320	6	A63607	A63607 Sequence 9
37	397	19.8	10320	6	AR281529	AR281529 Sequence
38	391.8	19.6	2654	5	FSCDYSTRO	M37645 Torpedo cal
39	387	19.3	387	6	AX538612	AX538612 Sequence
40	378.6	18.9	3161	10	MMGUTRPH	X83506 M.musculus
41	348	17.4	348	6	AX538611	AX538611 Sequence
42	330	16.5	333	6	AX538589	AX538589 Sequence
43	327	16.3	327	6	AX538588	AX538588 Sequence
44	324	16.2	324	6	AX538613	AX538613 Sequence
45	322.6	16.1	3499	9	HSU43519	U43519 Human dystro

ALIGNMENTS

RESULT 1
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

EATURES source Location/Qualifiers
1. .5462
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

RIGIN

Query Match 64.1%; Score 1283; DB 6; Length 5462;
Best Local Similarity 99.3%; Pred. No. 6.9e-311;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Y 691 GCTCCTGGACTGACCACACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGTTGACACAA 750
b 1547 GCTCCTGGACTGACCACACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGTTGACACAA 1606

Y 751 CCTGTGTTACTTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCITGATGTTG 810
b 1607 CCTGTGTTACTTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCITGATGTTG 1666

Y 811 GAGGTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAGTTTCTTGCC 870
b 1667 GAG-----CATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAGTTTCTTGCC 1717

Y 871 TGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGG 930
b 1718 TGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGG 1777

Y 931 CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTTCCAAGGT 990
b 1778 CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTTCCAAGGT 1837

Y 991 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAATAATCCTG 1050
b 1838 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAATAATCCTG 1897

Y 1051 AGATCCCTGGAAAGTCCGATGATGCAGTCCTGTACAAAGACGTTTGGATAACATGAAC 1110
b 1898 AGATCCCTGGAAAGTCCGATGATGCAGTCCTGTACAAAGACGTTTGGATAACATGAAC 1957

Y 1111 TTCAAGTGGAGTGAAGTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGT 1170
b 1958 TTCAAGTGGAGTGAAGTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGT 2017

Y 1171 TCTGACCAGTGAAGCGTCTGCACCTTTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG 1230
b 2018 TCTGACCAGTGAAGCGTCTGCACCTTTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG 2077

Y 1231 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCAGCAGCTCAGAAG 1290
b 2078 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCAGCAGCTCAGAAG 2137

Y 1291 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAGAACCTGTAATCATG 1350
b 2138 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAGAACCTGTAATCATG 2197

Y 1351 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 1410
b 2198 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 2257

Y 1411 CTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGATGTCACCTCGGCTT 1470
b 2258 CTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGATGTCACCTCGGCTT 2317

Y 1471 CTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCT 1530
b 2318 CTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCT 2377

Y 1531 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG 1590
b 2378 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG 2437

Y 1591 GATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTG 1650

Db 2438 GATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTG 2497

QY 1651 GCGGATCTCCTCAATCACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGA 1710
Db 2498 GCGGATCTCCTCAATCACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGA 2557

QY 1711 GAAATTGCGCCTCTGAAAAGAGAACGTTGAGGCACGTCATGACCTTGTCTGCCAGCTTACC 1770
Db 2558 GAAATTGCGCCTCTGAAAAGAGAACGTTGAGGCACGTCATGACCTTGTCTGCCAGCTTACC 2617

QY 1771 ACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGA 1830
Db 2618 ACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGA 2677

QY 1831 TGGAAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAGCAGCTGCGATGAAGCCACAGG 1890
Db 2678 TGGAAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAGCAGCTGCGATGAAGCCACAGG 2737

QY 1891 GACTTTGGTCCAGCATCTCAGCACTTCTTCTCCACGTCGTCCAGGGTCCCTGGGAGAGA 1950
Db 2738 GACTTTGGTCCAGCATCTCAGCACTTCTTCTCCACGTCGTCCAGGGTCCCTGGGAGAGA 2797

QY 1951 GCCATCTCGCCAAACAAAAGTGCCTACTATATCAACCACGAGACTCAAAACA 2001
Db 2798 GCCATCTCGCCAAACAAAAGTGCCTACTATATCAACCACGAGACTCAAAACA 2848

RESULT 2
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. .5952
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 5952;
Best Local Similarity 99.7%; Pred. No. 1.1e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
Db 2946 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 3005

QY 873 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCT 932
Db 3006 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCT 3065

QY 933 CTAGAAGACTCCAAGGGAGTAAAAGAGTGAACAATGGCAAGACCTCCAAGGTGA 992
Db 3066 CTAGAAGACTCCAAGGGAGTAAAAGAGTGAACAATGGCAAGACCTCCAAGGTGA 3125

QY 993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 1052
Db 3126 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 3185

QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCT 1112
Db 3186 ATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCT 3245

Y 1113 CAAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCTCATTTGGAAGCCAGTTC 1172
b |||||
3246 CAAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCTCATTTGGAAGCCAGTTC 3305
Y 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
b |||||
3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3365
Y 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTATGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 1292
b |||||
3366 AGATGATGAATTAAGCCGCGCAGGCACCTATTATGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 3425
Y 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
b |||||
3426 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 3485
Y 1353 TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAACT 1412
b TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAACT 3545
Y 1413 CTACCAGGAGCCCGAGAGCTGCTCTCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCT 1472
b CTACCAGGAGCCCGAGAGCTGCTCTCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCT 3605
Y 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAAACCTGCACTCCGCTGA 1532
b ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAAACCTGCACTCCGCTGA 3665
Y 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
b CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 3725
Y 1593 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
b TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 3785
Y 1653 CGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
b CGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3845
Y 1713 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
b AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905
Y 1773 TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATG 1832
b TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATG 3965
Y 1833 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAAGCCACAGGGA 1892
b GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAAGCCACAGGGA 4025
Y 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
b CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 4085
Y 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
b CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4134

ESULT 3
X114289
OCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
EFINITION Sequence 1 from Patent WO0129243.
CCESION AX114289
ERSION AX114289.1 GI:14031259
EYWORDS
OURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE 1

AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)
FEATURES Location/Qualifiers
source 1..5952
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"
ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 5952;
Best Local Similarity 99.7%; Pred. No. 1.1e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
Db |||||
2946 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 3005
QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db |||||
3006 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 3065
QY 933 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA 992
Db |||||
3066 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA 3125
QY 993 AATTGAAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAG 1052
Db |||||
3126 AATTGAAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAG 3185
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACCTT 1112
Db ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACCTT 3245
QY 1113 CAAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCAATTGGAAGCCAGTTC 1172
Db CAAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCAATTGGAAGCCAGTTC 3305
QY 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3365
QY 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTATGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 1292
Db AGATGATGAATTAAGCCGCGCAGGCACCTATTATGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 3425
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
Db GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 3485
QY 1353 TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 3545
QY 1413 CTACCAGGAGCCCGAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCT 1472
Db CTACCAGGAGCCCGAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCT 3605
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAAACCTGCACTCCGCTGA 1532
Db ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAAACCTGCACTCCGCTGA 3665
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
Db CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 3725


```

1593 TGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
1713 AATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
1833 GAAGCTTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCAATGAAGACCTGAAAGCCACAGGGA 1892
1953 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 1952
2073 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 2052
2193 CATCTCGCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 2201
2313 CATCTCGCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 2314

RESULT 4
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE
1 Chamberlain, J.S. and Harper, S.Q.
Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source 1..8689
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 8689;
Best Local Similarity 99.7%; Pred. No. 1.1e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1593 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCTG 872
1713 GGAAGAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCTG 3051
1833 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCT 932
2073 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCT 3111
2193 CCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGA 992
2313 CCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGA 3171
2433 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCCAAAAAATCCTGAG 1052
2553 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCCAAAAAATCCTGAG 3231
2673 ATCCCTGGAAAGGTTCCGATGATGACGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
```

```

Db 3232 ATCCCTGGAAAGGTTCCGATGATGACGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 3291
QY 1113 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTTCCTCCATTGGAAGCCAGTTC 1172
Db 3292 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTTCCTCCATTGGAAGCCAGTTC 3351
QY 1173 TGACCACTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 3352 TGACCACTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3411
QY 1233 AGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCGGACTTTCCAGCAGTTCCAGAACGA 1292
Db 3412 AGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCGGACTTTCCAGCAGTTCCAGAACGA 3471
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTATCATGAG 1352
Db 3472 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTATCATGAG 3531
QY 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 3532 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 3591
QY 1413 CTACCAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 3592 CTACCAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 3651
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAACCTGCACCTCCGCTGA 1532
Db 3652 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAACCTGCACCTCCGCTGA 3711
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGGA 1592
Db 3712 CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGGA 3771
QY 1593 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 3772 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 3831
QY 1653 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 3832 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3891
QY 1713 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATATGACCTTGTCTCGCCAGCTTACCAC 1772
Db 3892 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATATGACCTTGTCTCGCCAGCTTACCAC 3951
QY 1773 TTTGGGCATTGAGCTCTCACCCGATATAAAGCTCAGCAGCTCTGGAAGACCTGAACACAGATG 1832
Db 3952 TTTGGGCATTGAGCTCTCACCCGATATAAAGCTCAGCAGCTCTGGAAGACCTGAACACAGATG 4011
QY 1833 GAAGCTTCTGAGGTTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCACAGGGA 1892
Db 4012 GAAGCTTCTGAGGTTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCACAGGGA 4071
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTCCAGGGTCCCTGGGAGAGAGC 1952
Db 4072 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTCCAGGGTCCCTGGGAGAGAGC 4131
QY 1953 CATCTCGCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 2001
Db 4132 CATCTCGCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 4180

RESULT 5
AX538624 LOCUS AX538624 11443 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 44 from Patent WO0229056.
ACCESSION AX538624
VERSION AX538624.1 GI:25271175
KEYWORDS synthetic construct
SOURCE synthetic construct
```

artificial sequences.

REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES Location/Qualifiers
1. .11443
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

RIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 11443;
Best Local Similarity 99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Y 813 GGTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTG 872
b 5746 GGAAGAACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTG 5805

Y 873 GCTTACAGAACTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAGGCT 932
b 5806 GCTTACAGAACTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAGGCT 5865

Y 933 CCTAGAAGACTCCTAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGGTGA 992
b 5866 CCTAGAAGACTCCTAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGGTGA 5925

Y 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAATCCTGAG 1052
b 5926 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAATCCTGAG 5985

Y 1053 ATCCCTGGAAGTTCCGATGATGCGAGTCTGTTCACAAAGACGTTTGGATAACATGAACACTT 1112
b 5986 ATCCCTGGAAGTTCCGATGATGCGAGTCTGTTCACAAAGACGTTTGGATAACATGAACACTT 6045

Y 1113 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTTCCTCAATTTGGAAGCCAGTTTC 1172
b 6046 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTTCCTCAATTTGGAAGCCAGTTTC 6105

Y 1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
b 6106 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 6165

Y 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA 1292
b 6166 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA 6225

Y 1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAGAACCTGTATCATGAG 1352
b 6226 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAGAACCTGTATCATGAG 6285

Y 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
b 6286 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 6345

Y 1413 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
b 6346 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 6405

Y 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCATCCGCTGA 1532
b 6406 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCATCCGCTGA 6465

Y 1533 CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGAACTTCAAGAGGCCACGGA 1592
b 6466 CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGAACTTCAAGAGGCCACGGA 6525

Y 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGG 1652
b 6526 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGG 6585

QY 1653 CGATCTCTCATTTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGA 1712
Db 6586 CGATCTCTCATTTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGA 6645

QY 1713 AATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCATTAATGACCTTGCTCGCCAGCTTACCAC 1772
Db 6646 AATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCATTAATGACCTTGCTCGCCAGCTTACCAC 6705

QY 1773 TTTGGGCATTGAGCTCTCACCGTATAAACCCTCAGCACTCTTGAAAGACCTGAAACACCAAGATG 1832
Db 6706 TTTGGGCATTGAGCTCTCACCGTATAAACCCTCAGCACTCTTGAAAGACCTGAAACACCAAGATG 6765

QY 1833 GAAGCTTCTGAGGTGGCGTCTGAGGACCGAGTCAGGCAGCTGCTCATGAAGCCACACAGGA 1892
Db 6766 GAAGCTTCTGAGGTGGCGTCTGAGGACCGAGTCAGGCAGCTGCTCATGAAGCCACACAGGA 6825

QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGC 1952
Db 6826 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGC 6885

QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
Db 6886 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 6934

RESULT 6
AX538627
LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
1. .12057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN.

Query Match 59.1%; Score 1182.6; DB 6; Length 12057;
Best Local Similarity 99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTG 872
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTG 8319

QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db 8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
Db 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAATCCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAATCCTGAG 8499

QY 1053 ATCCCTGGAAGGTTCCGATGATGCGCTCTGTTCACAAAGACGTTTGGATAACATGAACCTT 1112
Db 8500 ATCCCTGGAAGGTTCCGATGATGCGCTCTGTTCACAAAGACGTTTGGATAACATGAACCTT 8559

2Y	1113	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTC	1172
2b	8560	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTC	8619
2Y	1173	TGACCACTGGGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA	1232
2b	8620	TGACCACTGGGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA	8679
2Y	1233	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGACCTTTCCAGCAGTTCAGAAAGCA	1292
2b	8680	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGACCTTTCCAGCAGTTCAGAAAGCA	8739
2Y	1293	GAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTAATCATGAG	1352
2b	8740	GAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTAATCATGAG	8799
2Y	1353	TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTTGGAAGGACTAGAGAAACT	1412
2b	8800	TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTTGGAAGGACTAGAGAAACT	8859
2Y	1413	CTACCAGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCCGAAGATGTCACCTCGGCTTCT	1472
2b	8860	CTACCAGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCCGAAGATGTCACCTCGGCTTCT	8919
2Y	1473	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA	1532
2b	8920	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA	8979
2Y	1533	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA	1592
2b	8980	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA	9039
2Y	1593	TGAGCTGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGG	1652
2b	9040	TGAGCTGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGG	9099
2Y	1653	CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA	1712
2b	9100	CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA	9159
2Y	1713	AATTGCGCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	1772
2b	9160	AATTGCGCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	9219
2Y	1773	TTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCAGATG	1832
2b	9220	TTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCAGATG	9279
2Y	1833	GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGCAGCTGCATGAAGCCCCACAGGGA	1892
2b	9280	GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGCAGCTGCATGAAGCCCCACAGGGA	9339
2Y	1893	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC	1952
2b	9340	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC	9399
2Y	1953	CATCTCGCAAAACAAAGTCCCTACTATATCAACCCAGGAGACTCAAACA	2001
2b	9400	CATCTCGCAAAACAAAGTCCCTACTATATCAACCCAGGAGACTCAAACA	9448
RESULT 7			
HSMDR	Human mRNA for dystrophin.	12446 bp	mRNA
LOCUS	X14298		
DEFINITION	Dmd gene; Duchenne muscular dystrophy; dystrophin.		
ACCESSION	X14298.1	GI:30845	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 12446)		

AUTHORS	Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.
TITLE	Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL	Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE	89345106
PUBMED	2668885
REFERENCE	2 (bases 1 to 12446)
AUTHORS	Rosenthal,A.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR
COMMENT	see also M18533 and M20250 for Dmd segs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES	Location/Qualifiers
source	1. .12446
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="X chromosomal, Xp21."
	/tissue_type="muscle"
	/dev_stage="fetal and adult."
	99. .11156
CDS	/note="unnamed protein product; dystrophin (AA 1 - 3685)"
	/codon_start=1
	/protein_id="CAA32479.1"
	/db_xref="GI:30846"
	/db_xref="GOA:P11532"
	/db_xref="SWISS-PROT:P11532"
	/translation="MLWEEVEDCYEREDVOKKFTTKWVNAQFSKFGKHENLFSDL QDGRLLDLLGLTGQKLPKEKGSTRVHALNVNNAKALRVLQNNVLDLVNIGSTDIVDG NKKLTGLIWNIIILHWQVNMKNIMAGLOPTNSEKILLSWVRQSTPNYPQVNVINFT TSWSDGLALNALIHSRPPDLFDWNSVVCQSATORLEHAFNIARYQLGIEKLDPEDV DTTYDKKSILMYITSLFQVLPQQVSIEAIOEVEMLPRPPKVTKEEHFQLHHQMHYSQ QITVSLAQGYERTSSPKRPFKSYAYTAAVTTSDPRTSPFPSSHLEAPEDKSFSSSL MESEVNLDRYOTALSEVLWSLAEADTLQAOGEISNDVEVVKDOFHTHEGYMMDLTAH QGRVNIQLGSKLIGTKLSEDEETEVOQMNLNLSRWECRLVASMESQSNLHRVLM DLQNKLKELNDWLTKEERTKMEEEPLGPDLEDLKRQVQKHVLOEDLEQEQVRVN SLTHMVVVDESSEGDHATAALEEQKVLGDRWANI CRWTEDRVWLLQDILLKQRLTE EQCLFSAWLSEKEDAVNKIHTTGFKDQNMELSSLOKLAVLKADLEKKQSMGKLYSIK QDLSLTLKNKSVTKTEAWLDNFARCDNLVQKLEKSTAQISOAVTTTPSLTQTVM ETVTVTTREQILVHAQEELPPPPQKKRQITVDSEIRKRLVDITELHWSITRSEA VLOSPEFAIFRKEGFSDLKEKVNAIERAEKFRKLQDASRGOALVEQMNVGNVA DSIQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQLEQMTTAAENWLKIQ PTPPSEPTAIKSQLKICKDEVNRLSGLOIERLKIQSIALKEKGQPMFLDADFVAF TNHFQVFSVDQAREKELQITFDTLPPMYQETMSAIRTWVQSQSETKLSIPQLSVTDY EIMEQRLGELQALQSLQEQSGLYLSTTVKEMSKKAPSEISRKYSQSEFEIEGRWK KLSQLVEHCQKLEBQMNKLRIQNHIIQLKKWMAEVDVFLKEWPALGDSIILKKQL KQCRLLVSDIQTIOPSLNSVNEGQKINAEAPFASRLTELKELTQWDMHCQQVY ARKEALKGLEKTVSLQKDLSEMHWTQABEYLERDFEYKTDDELQKAFEEKRAK BEAQKEAKVLLTESVNSVIAQAPPVAQBALKELETLTTNYQWLCRLNGKCKTLE EVWACHWELLSYLEKANKWLNVEFEKLTENIPGAAEEISEVLDLENLMRSHEDNP NOIRILAQTLTDGGVMDLNEELETFNSRWRELHEEAARRQKLLQSIQSAQETENS LHIIQESLTFIDKQLAAIADKVDAQMPQEAQKIQSDLTSHEISLEEMKXNQKEA AQVLSQIDVAQKQLQDVSMKFRLLFQKPANFEQRLQESKMILDEVKMHLPALETKSV ENEVGSQLNHCVNLYKSLSEVKSEVMVKTGRQIVQKKQNTENPKELDERVATLKLHY NELGAKVTERKQKLEKCLKSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE VAWGKATQKEIEKQKVLKSI TEVGEALTVLGKKTLEVEDKLSLNSNWIATVSRAE EWLNLLEYQKHMETFDQNVDHITKWIIOADTLLDESEKKPQOKEDVLKRLKAELND TRPKVDSTRDQAANLMANHGDCRKLVEPOISELNHRFAAISRIKTKGKASIPLKELE QFNSDIQKLLPLEBAIQGGVNLKEEDFNKDMNEDNEGTVKELLQORDNLQORITDER KSEBEIKIQQLLOTXNALKDLRSQRRKKALEISHQWYQYKRDADLLKCLDDIEKKL ASLPEPRDERKIKEIDRELOKKEELNAVRRQAEGLSDGAAMAVEPTQIOLSKRWRE TBSKPAQFRRLNFAQIHTVREEMVMVMTEDMPLSISVPSTYLTETHVSQALLEVEQ LINAAPDLCAKDFEDLFKQEEESLKNIKDSLQSSSGRIDIIHSKTAALQSATPVERVKL QEALSQDLDFQWEKVNMYKDRQFRDREVEKWRFRHYDIKIFNWLTEAEQFLRKTQI PENWEHAKYKWLKELQDGIQGRQTVVRLNATGEEIIQSSKTDASILOEKLGSNLN RWQEVCKQLSDRKIKLEEQKNILSEFQRLNEFVLEWEEADNIASIPLEPGKEQQLKE KLEQVKLLVEELPLRQGIKQLNETGGPVLVSAPISPEEQDKLENKCLKQTNLQWIKVS RALPEKQGEIEAQIKDLQGLEKLEDEQLNHLLLWLSPIRNOLEIYNQPNQEGPFD VKETEIAVQAKQPDVVEILSKGQHLKYKEKATQPVKRLKLELSDSEWKAVNRLQLERA

KOPDLAPGLTTIGASPTOTVTLVTPQVVTKETATSKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSORVMVGDLEDINEMI IKQXATMQDLEQRRQLBELITAAQNLK
NKTSNQEARTIITDRIERIQNWDEVOEHLQNRQQLNEMVKDSTOWLEAKEEAEQVL
GQARAKLESWKEGPTYVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
DTRKVMITENINASWRSIHKRVSEREAALTEHRLLOQFPLDLEKFLAWLTEAETTA
NVLDQATRKERLLEDKSGVKELMKQWQDLOGEIHAETHDVYHNLNDENSQKILRSLEGS
DAVLLQRRLLDNMFKWSERKKSNIIRSHLEASSDQWKRLHLSDLELLVWLQKXDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTEKPEVIMSTLETVRIFLIEQPLEGLEKYQ
EPRELPPPEERAOVTRLLRKOAEVNTWEKLNLSADWOKIDEHLERLOELOEAT
ELDLKLRQAEVIKGSWPQVGDLLIDSLQDHLKVKYALRGEIAPLKENVSHVNDLARQL
TTLGIQSPYNLSTLEDLNRWKLQVAVEDRVRLHEAHRDFGSPASQHFLLSTSVQGP
WERAI SPNKVPYYINHETQTCWDHPKMTLYQSLADLNNRVFSYRTAMKLLRRLQKA
LCLDLLSLSAACDALDQNLKQNDQPMDDIQLINCLTTIYDRLEQHNHNLVNVPLCVD
MCLNWLNNVYDTRGRIRVLSFKGTIISLCAHLEDKRYLYLFQVASSTFCDQRRLL
GLLLHDSIQIPROLGEVASFGGSNI EPSVRSQFANNKPEIEAALFLDMRLEPQSM
VWLPVYLRHVAATAKHQAKNICKECP IIGERYRSLKHFNHYDCOSCFFSGRVARGH
KMHPVMEYCTPTTSGEDVDFAKVLKNKFKTKRYFAKHPRMGYLPVQTVLEGDMET
PVTLLNFWPVD SAPASSPOLSHDDTHSRIEYASRLAEMENSNGSYLNDISIPNESID
DEHLLIQHYCQSLNQDPSLQSPRSPAQILISLESEERGELERILADLEENRNLAQY
DRKQOHEHKGSLPLSPSPPEMMPTSPQSPRDAELIAEAKLLRQHRLEARMQILEDH
NKQLESQHLRLLQLEQPAEAKVNGTTVSSPSTSLQRSDDSSQPMLLRVVGSQTSDSM
GEEDLLSPQDSTSTGLEEYMEQLNNSFPSSRGENTPGKPMREDTY"

RIGIN

Query Match		59.1%;	Score 1182.6;	DB 9;	Length 12446;	
Best Local Similarity		99.7%;	Pred. No. 1.2e-285;			
Matches 1185;		Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
/	813	GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCCTG	872			
/	8150	GGAAGAACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCCTG	8209			
/	873	GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT	932			
/	8210	GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT	8269			
/	933	CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACTCTCAAGGCTGA	992			
/	8270	CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACTCTCAAGGCTGA	8329			
/	993	AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAATAATCCTGAG	1052			
/	8330	AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAATAATCCTGAG	8389			
/	1053	ATCCCTGGAGGTTCCGATGATGACGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT	1112			
/	8390	ATCCCTGGAGGTTCCGATGATGACGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT	8449			
/	1113	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTTC	1172			
b	8450	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTTC	8509			
/	1173	TGACCAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	1232			
b	8510	TGACCAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	8569			
/	1233	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCCAGAAGCA	1292			
b	8570	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCCAGAAGCA	8629			
/	1293	GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG	1352			
b	8630	GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG	8689			
/	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT	1412			
b	8690	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT	8749			
/	1413	CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTACTCGGCTTCT	1472			
b	8750	CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTACTCGGCTTCT	8809			
/	1473	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCTCGCTGA	1532			

Db	8810	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGAAAAAATTGAACCTGCACCTCCGCTGA	8869			
QY	1533	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA	1592			
Db	8870	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA	8929			
QY	1593	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG	1652			
Db	8930	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG	8989			
QY	1653	CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA	1712			
Db	8990	CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA	9049			
QY	1713	AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	1772			
Db	9050	AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	9109			
QY	1773	TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG	1832			
Db	9110	TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG	9169			
QY	1833	GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGA	1892			
Db	9170	GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGA	9229			
QY	1893	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC	1952			
Db	9230	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC	9289			
QY	1953	CATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA	2001			
Db	9290	CATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA	9338			
RESULT 8						
AX409637						
LOCUS		AX409637	13957 bp	DNA	linear	PAT 14-JUN-2002
DEFINITION		Sequence 2284 from Patent WO0229103.				
ACCESSION		AX409637				
VERSION		AX409637.1 GI:21442342				
KEYWORDS		Homo sapiens (human)				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens				
REFERENCE		1				
AUTHORS		Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.				
TITLE		Gene expression profiles in liver cancer				
JOURNAL		Patent: WO 0229103-A 2284 11-APR-2002;				
FEATURES		GENE LOGIC INC (US)				
source		Location/Qualifiers				
		1. .13957				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
		/note="EMBL/GenBank Accession No. M18533"				
ORIGIN						
Query Match		59.1%;	Score 1182.6;	DB 6;	Length 13957;	
Best Local Similarity		99.7%;	Pred. No. 1.2e-285;			
Matches 1185;		Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
QY	813	GGTACCTACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAGTTTCTTGCTG	872			
Db	8260	GGAAGAACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAGTTTCTTGCTG	8319			
QY	873	GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCT	932			
Db	8320	GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCT	8379			
QY	933	CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGA	992			

8380 CCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 8439
993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG 1052
8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG 8499
1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 8559
1113 CAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAAGTCCCATTTGGAGCCAGTTC 1172
8560 CAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAAGTCCCATTTGGAGCCAGTTC 8619
1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
8620 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 8679
1233 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCA 1292
8680 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCA 8739
1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
8740 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 8799
1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAACT 1412
8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAACT 8859
1413 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCT 1472
8860 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCT 8919
1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTCAGAACCTTCAAGAGCCACGGA 1532
8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTCAGAACCTTCAAGAGCCACGGA 8979
1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACCTCCAGGAACCTTCAAGAGCCACGGA 1592
8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACCTCCAGGAACCTTCAAGAGCCACGGA 9039
1593 TGAGCTGGACCTCAAGCTGCGCCAGTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 1652
9040 TGAGCTGGACCTCAAGCTGCGCCAGTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 9099
1653 CGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
9100 CGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
1713 AATTGCGCCTCTGAAAAGAGAACGTTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1772
9160 AATTGCGCCTCTGAAAAGAGAACGTTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 9219
1773 TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 1832
9220 TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 9279
1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAAGCCACAGGGA 1892
9280 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAAGCCACAGGGA 9339
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGAGAGAGC 1952
9340 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGAGAGAGC 9399
1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 2001
9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 9448

AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 13957;
Best Local Similarity 99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCTG 872
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCTG 8319
QY 873 GCTTACAGAAAGCTGAAAAACAACCTGCCAATGTCTCTACAGGATGCTTACCCGTAAGGAAAGCT 932
Db 8320 GCTTACAGAAAGCTGAAAAACAACCTGCCAATGTCTCTACAGGATGCTTACCCGTAAGGAAAGCT 8379
QY 933 CTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 992
Db 8380 CTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 8439
QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCT 1112
Db 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCT 8559
QY 1113 CAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTC 1172
Db 8560 CAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTC 8619
QY 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCA 1292
Db 8680 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCA 8739
QY 1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
Db 8740 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 8799
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAACT 8859
QY 1413 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCT 1472
Db 8860 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTCAGAACCTTCAAGAGCCACGGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTCAGAACCTTCAAGAGCCACGGA 8979

1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGGA 1592
|||||
8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGGA 9039
|||||
1593 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
|||||
9040 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9099
|||||
1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
|||||
9100 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
|||||
1713 AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1772
|||||
9160 AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 9219
|||||
1773 TTTGGGCATTGAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832
|||||
9220 TTTGGGCATTGAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 9279
|||||
1833 GAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCTGAAGCCCACAGGGA 1892
|||||
9280 GAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCTGAAGCCCACAGGGA 9339
|||||
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCTGGGAGAGAGC 1952
|||||
9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCTGGGAGAGAGC 9399
|||||
1953 CATCTGCGCAAAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAACA 2001
|||||
9400 CATCTGCGCAAAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAACA 9448
|||||

RESULT 10
UMDYS
OCUS HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
EFINITION Homo sapiens dystrophin (DMD) mRNA, complete cds.
CCESION M18533 M17154 M18026 M20250
ERSION M18533.1 GI:181856
EYWORDS
OURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1699)
AUTHORS Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.
TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
JOURNAL Cell 50 (3), 509-517 (1987)
MEDLINE 87273512
PUBMED 3607877
REFERENCE 2 (bases 1678 to 3830)
AUTHORS Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3 (bases 1 to 13957)
AUTHORS Koenig,M., Monaco,A.P. and Kunkel,L.M.
TITLE The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein
JOURNAL Cell 53 (2), 219-226 (1988)
MEDLINE 88194521
PUBMED 3282674
COMMENT On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168,

1691-1810, and 1169-3011.
Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AAA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFKQKHENLFSDL
QDGRRLDLLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNVLDVNI GSTDIVDG
NHKLTGLIWNIIHLHVQVKNVMKNI MAGLQQTNSKILLSWRQSTRNYPQVNVINFT
TSWSDGLALNALIHSRPPDLFDWNSVVCQASATQRL EHAFNARIYQLGIEKLLDPEDV
DTTPDKKXILMYITSLFQVLPQVSI EAIQEVEMLP RP PKVTK EEFHQLHHQMHYSQ
QITVSLAOGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSSHLEAPEDKSFSSSL
MESEVNLDRYQTAL EEEVLSWLLSAEDTLQA GEI SNDEVVKDQFHTHEGYMMDLTAH
QGRVGNILQLSGKLTGKLTSEDETEVEGEQMNLLNSRWECLRVASMEKQSNLHRVLM
DLQNKQLKELNDLTKTEERTRKMESEPLGDL EDKRQVQKHVKLQEDLEQEQRVN
SLTHMVVVDSSGDHATAALEEQVLKVLGDRWANI CRWTEDRWLLQDILLKWORLTE
EQCLFSAWLSEKEDAVNKIHTTGFKDONEMLSLQKLA VLKADLEKKKQSMGKLYSLK
QDLLSTLKNKSVTKTEAWLDFNARCWNLVQKLEKSTAQISQAVTTTPSLTQTVM
ETVTTVTREQILVKHAQEELPPPPPKKRCQITVDSEIRKRLDVDTLHSWITRSEA
VLQSPFAIFRKEGNFSDLKEKVNAIEREAK EFRKLQDASRSQAALVEQMVNEGVA
DSIKOASEQLNSRWIEFCOLL SERLNLWEYQNNIIAFYNOLQQL EQMTTTAENWLKIQ
PTTPSEPTAISKQKICKDEVNLSGLQPIERLKIQSI ALKEKGQGMFLDADFVAF
TNHFQVFSVDVQAREKELQTI FDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
EIMEQRLGELQALQSSLOEQQSGLYLSTTVKEMSKKAPSEISRKYQSEFEIEGRWK
KLSSQLVEHCQKLEEQMNKLRKIQNHIIQTLKKWMAEVDVFLKEEWPALGDSEILKKQL
KQCRLLVSDIIQTTQPSLNSVNEGQKTKNEAEPEFASRLTELKELNTQWDMHCQQVY
ARKEALKGGLEKTVSLQKDLSEHWMVTA EEEYLERDFEYKTPDELQKAVEEMKRAK
EEAQKKEAKVKLTESVNSVIAQAPVQA EALKLELTLTNNQWLCTRLNGCKRTLE
EVAWCWHELLSYLEKANKWLN EVEFKLKTENIPGGAEEISEVLDLSENLMRHS EDNP
NQIRILAQTLTDGGVMD ELINEELETENSRRELHEEAVRRQKLEQSIQSAQET EKS
LHLIQESLTFIDKQLAAYIADKVDA AQPQEAQKIQSDLTSH EISLEEMKKHQGKEA
AQRVLSQIDVAQKLDVSMKFR LFKQPANFELRQKEDVLRKLKAE LND
QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKOTENPKELDERVTALKLHY
NELGAKVTERKQOLEKCLKLSRKRMKEMNLTEWLAATMELTKRS AVEGMPSNL DSE
VANGKATQKEIEKQKVLKSI TEVGEALKTVLGKETLVEDKLSLLNSNVIATSRAE
EWNLNLL EYQKHMTFDQNV DHI TWIIQADTL DLESEKKKPOQKEDVLRKLKAE LND
IRPKVDSTRDQAANLMANRGDHCRLVEPQISELNHRFAAISHRIKTGASIP LKLE
QFNSDIOKLLEPLEABIQGVNLKEEDFNKM DNEGTVEKLIQRGDNLQORITDER
KREEIKIKQQLLQTKHNALKD LRSQRKKALEISHQWYQYKRAQDDLKCLDDIEKKL
ASLPEPRDERKIKIDRELQKK EELNAVRRAEGLSEGDGAAMAVEPTQIOLSKRWRE
IESKFAQFRRLNFAQIHTVREETMVMTEDMPLEISYVPSTYLTETHVSQALLEVEQ
LLNAPDLCAKDFEDL FKQEE SLKNIKDSLQOSSGRIDIIH SKTAALQSATPVERVKL
QEALSQLDFQWEKVNMYKDRQRFDRSVEKWRFRHYDIKIFNQLWTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGQROT VVTRLNATGEEIIQQSKTIDASILQEKLGSLNL
RWQEVCKQLSDRKKRLEEQNILSEFORDLNEFVLWLEADNIASIPLEPGKEQQLKE
KLEQVKLLVEELPLRQGI LKQLNETGGPVLVSAPISPEEQDKLENKIKQTNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKKLEDEEQNLHLLLLWLSPIRNQLEIYNQPNQEGPFD
VQET EIAVQAQKQPDVEEILSKQHLYKKEKPATQPVKRLDLSSEKWKAVNRLLOELRA
KQPD LAPGLTTIGASPTQT VTLVTPVVTKETAI SKLEMPSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSQRVMVGDLEINEMI IKQKATMQDLEORRPQLEELITAAQN LK
NKT SNQEARTIITDRIERI QNWDEVQEH LQNRRLQQLNEMLDKSTQWLEAKEEAEQVL
GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKL RDYSAD
DTRKVMITENINASWRSIHKRV SEREALEETHRL LQQFPDLEKFLAWL TEAETTA
NVLQDATRKERLLED SKGVKELMKQWQDLQGEIEAHTDVVHMLDENSQKILRSLEGS D
DAVLLQRRLDNNMFKWSELRS LNIRHLEASSDQWKRLHSLQELLLVWLQKDDDEL
SRQAPIGGDFPAVQXQNDVHRAF KRELKTKPEVMSTLETVRI FLTEQPLEGKLYQ
EPRELPP EERAQNVTRL LRKQAE EVNTEWELKNLHSDWQRKIDETLERLQELQEATD
ELDLKLROAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTIGIQLSPYNLSTLDLNTRWKLLQVAVEDRVQLHEHRDFGPASQHFSTSVQGP
WERAI SPKNKVPYIYNHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA

LCDLLLSAACDALDQHNKQNDQPMDIQIINCLTTIYDRLEQERNNLVNVPLCVD
MCLNWLNVYDTRGTRIRVLSEFKTGIISLCKAHLEDKYRYLFXQVASSSTGFCQDORRL
GLLHDSIQIPROLGEVASFGSNIEPSVRSCTFANNKPEIEAALFLDWMRLPQSM
VWLPVLHRAVAEAETAKHQAKNICKECPIIGFRVSLKHFNVDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSSGEDVRDEFAKVLKNKFRTRYPAKHPRMGYPVQTVLEGNMET
PVTLLNFWPVD SAPASSPOLSHDDTHSRIEHYASRLAEMENSNGSYLNDISIPNESID
DEHLIQHYCQSLNQDSPLOSPSPAQILISLESBERGELERILADSEERNLNQAEY
DRLKQHEHKGLSPLPSPENMPTSPQSPRDAELIABAKLLRQHKGLRQLEARMQILEDH
NKQESQLHRLRQLLEQPOBAKVNGTTVSSPSTSLQRSOSSQPMLELVRVGSQTSDSM
GEEDDLSPQDTSSTGLEEVMEQLNNSFPSSSRGNTPGKPMREDTM"

ORIGIN

Query Match 59.1%; Score 1182.6; DB 9; Length 13957;
Best Local Similarity 99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
DB 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 8319

QY 873 GCTTACAGAAGCTGAACAACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCT 932
DB 8320 GCTTACAGAAGCTGAACAACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCT 8379

QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
DB 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAAAANTCCTGAG 1052
DB 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAAAANTCCTGAG 8499

QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTGTACAAAGACGTTTGGATAACATGAACCTT 1112
DB 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTGTACAAAGACGTTTGGATAACATGAACCTT 8559

QY 1113 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTT 1172
DB 8560 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTTT 8619

QY 1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
DB 8620 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 8679

QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTCCAGCAGTTCCAGAAGCA 1292
DB 8680 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTCCAGCAGTTCCAGAAGCA 8739

QY 1293 GAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAGAAACCTGTAATCATGAG 1352
DB 8740 GAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAGAAACCTGTAATCATGAG 8799

QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACT 1412
DB 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACT 8859

QY 1413 CTACCAGGAGCCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCCAAGATGTCACTCGGCTTCT 1472
DB 8860 CTACCAGGAGCCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCCAAGATGTCACTCGGCTTCT 8919

QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTTCAAGAGCCGCTGA 1532
DB 8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTTCAAGAGCCGCTGA 8979

QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
DB 8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039

QY 1593 TGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
DB 9040 TGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9099

QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712

DB 9100 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
QY 1713 AATTGCGCCTCTGAAAGAGAAACGTTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCAC 1772
DB 9160 AATTGCGCCTCTGAAAGAGAAACGTTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCAC 9219
QY 1773 TTTGGGCATTTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
DB 9220 TTTGGGCATTTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATG 9279

QY 1833 GAAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
DB 9280 GAAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 9339

QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGTCCTTGGGAGAGAGC 1952
DB 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGTCCTTGGGAGAGAGC 9399

QY 1953 CATCTGCCAAACAAAGAGTGCCTACTATATCAACCACGAGACTCAAAACA 2001
DB 9400 CATCTGCCAAACAAAGAGTGCCTACTATATCAACCACGAGACTCAAAACA 9448

RESULT 11
AR220819
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source 1..13977
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 13977;
Best Local Similarity 99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
DB 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 8319

QY 873 GCTTACAGAAGCTGAACAACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCT 932
DB 8320 GCTTACAGAAGCTGAACAACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCT 8379

QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
DB 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAAATCCTGAG 1052
DB 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAAATCCTGAG 8499

QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTCACAAAGACGTTTGGATAACATGAACCTT 1112
DB 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTCACAAAGACGTTTGGATAACATGAACCTT 8559

QY 1113 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTT 1172
DB 8560 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTT 8619

QY 1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232

b 8620 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCTGAA 8679
y 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTGAGGCGCAGCTTTCAGCAGCTTCAGAAAGCA 1292
b 8680 AGATGATGAATTAAGCCGCGCAGGCACCTATTGAGGCGCAGCTTTCAGCAGCTTCAGAAAGCA 8739
y 1293 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
b 8740 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 8799
y 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAAACT 1412
b 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAAACT 8859
y 1413 CTACCAAGGAGCCAGAGAGCTGCCTCCTGAGAGAGAGCCAGAAATGAACTGCTCGGCTCT 1472
b 8860 CTACCAAGGAGCCAGAGAGCTGCCTCCTGAGAGAGAGCCAGAAATGAACTGCTCGGCTCT 8919
y 1473 ACGAAAGCAGGCTGAGAGGCTCAATACTGAGTGGGAAAAATGAACTGCTCGGCTCT 1532
b 8920 ACGAAAGCAGGCTGAGAGGCTCAATACTGAGTGGGAAAAATGAACTGCTCGGCTCT 8979
y 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
b 8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 9039
y 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGG 1652
b 9040 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGG 9099
y 1653 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
b 9100 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
y 1713 AATTGCGCTCTGAAAGAGAAAGTGAAGCACTGAGTCAAGGGATCCTGGCAGCCCGTGG 1772
b 9160 AATTGCGCTCTGAAAGAGAAAGTGAAGCACTGAGTCAAGGGATCCTGGCAGCCCGTGG 9219
y 1773 TTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACCCAGATG 1832
b 9220 TTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACCCAGATG 9279
y 1833 GAAGCTTCTGCAGGTGGCGTGCAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 1892
b 9280 GAAGCTTCTGCAGGTGGCGTGCAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 9339
y 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGTCCTTGGGAGAGAGC 1952
b 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGTCCTTGGGAGAGAGC 9399
y 1953 CATCTCGCCAAACAAAGTGCCTACTATATCAACCCAGAGACTCAAAACA 2001
b 9400 CATCTCGCCAAACAAAGTGCCTACTATATCAACCCAGAGACTCAAAACA 9448

ESULT 12
X538620
OCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
EFINITION Sequence 40 from Patent WO229056.
CCENSION AX538620
ERSION AX538620.1 GI:25271166
EYWORDS synthetic construct
OURCE synthetic construct
ORGANISM artificial sequences.
EFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES Location/Qualifiers
source 1. 5339

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

Query Match 59.0%; Score 1180; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 5e-285;
Matches 1180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 822 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTCCTTCCCTGGCTTACAGA 881
Db 1546 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTCCTTCCCTGGCTTACAGA 1605
QY 882 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 941
Db 1606 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1665
QY 942 CTCCAAGGAGTAAAGAGCTGATGAAACAATGSCAAGACCTCCAAGGTGAATTTGAAGC 1001
Db 1666 CTCCAAGGAGTAAAGAGCTGATGAAACAATGSCAAGACCTCCAAGGTGAATTTGAAGC 1725
QY 1002 TCACACAGATGTTTATCACAACCTGGATGAAACACAGCCAAAAATCCTGAGATCCCTGGA 1061
Db 1726 TCACACAGATGTTTATCACAACCTGGATGAAACACAGCCAAAAATCCTGAGATCCCTGGA 1785
QY 1062 AGGTTCCGATGATGCGAGTCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1121
Db 1786 AGGTTCCGATGATGCGAGTCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1845
QY 1122 TGAACCTCGGAAAAAGTCTCTCAACATTAGTCCATTTGGAAGCCAGTCTTGACCAGTG 1181
Db 1846 TGAACCTCGGAAAAAGTCTCTCAACATTAGTCCATTTGGAAGCCAGTCTTGACCAGTG 1905
QY 1182 GAAGCTCTGCACCTTTCTCTGCAGGAACTTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1241
Db 1906 GAAGCTCTGCACCTTTCTCTGCAGGAACTTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1965
QY 1242 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGATGT 1301
Db 1966 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGATGT 2025
QY 1302 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1361
Db 2026 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2085
QY 1362 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1421
Db 2086 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 2145
QY 1422 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGAAATTTGAACCTCGCTGACTGGCAGAG 1481
Db 2146 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGAAATTTGAACCTCGCTGACTGGCAGAG 2205
QY 1482 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTCGCTGACTGGCAGAG 1541
Db 2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTCGCTGACTGGCAGAG 2265
QY 1542 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1601
Db 2266 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 2325
QY 1602 CCTCAAGCTGCGCCAAAGTGAAGTATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCT 1661
Db 2326 CCTCAAGCTGCGCCAAAGTGAAGTATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCT 2385
QY 1662 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCC 1721
Db 2386 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCC 2445
QY 1722 TCTGAAAGAGAAAGTGAAGCAGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 1781
Db 2446 TCTGAAAGAGAAAGTGAAGCAGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2505

QY	1782	TCAGCTCTCACCGTATAA	CCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCT	1841
Db	2506	TCAGCTCTCACCGTATAA	CCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCT	2565
QY	1842	GCAGGTGGCCGTCGAGGACCGAGT	CAGGCAGCTGCATGAAGCCCAACAGGGACCTTGGTCC	1901
Db	2565	GCAGGTGGCCGTCGAGGACCGAGT	CAGGCAGCTGCATGAAGCCCAACAGGGACCTTGGTCC	2625
QY	1902	AGCATCTCAGCACTTTCTTT	CCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	1961
Db	2626	AGCATCTCAGCACTTTCTTT	CCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	2685
QY	1962	AAACAAAGTGCCCTACTATAT	CAACCAAGAGACTCAACA	2001
Db	2686	AAACAAAGTGCCCTACTATAT	CAACCAAGAGACTCAACA	2725

RESULT 13
 AF070485
 LOCUS 13887 bp mRNA linear MAM 09-DEC-1998
 DEFINITION Canis familiaris dystrophin mRNA, complete cds.
 ACCESSION AF070485
 VERSION AF070485.1 GI:3982750
 KEYWORDS
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 13887)
 Carville, K.S., Mann, C.J., Schatzberg, S.J. and Wilton, S.D.
 Direct Submission
 Submitted (04-JUN-1998) ANRI, Pathology, University of Western
 Australia, Verdun Street, Nedlands, WA 6018, Australia
 FEATURES
 source
 1. 13887
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /strain="breed Golden retriever"
 /db_xref="taxon:9615"

/translation="MLWEEVEEDCYEREDVQKKTFFKWVNAQFSKFGKHIEIENLFSDI
 QDGRRLDLLLEGLTGQKLPKEGSTRVHALNNVNKALRVLQKNVDLVNIGSTDIVDG
 NNHKLTLGLTWNILHMQVKVNMIMAGLOQTNSEKILLSWVRQSTRNPQVNVINFT
 TSWSDGLKALNALIHSHRPDLFDWNSVQCOQASTORLEHAFNAIKYQGLIEKLLDPEDV
 ATTPYDGLKALIMYITSLFQVLPOOVSIETAEQVEMLPSPQVSTRREEHFOFHQMHSYQ
 QITVSLAQGVERAPSPFKPRFKSYATQAAVYVTTSDTRSPLPQSHLETPEDKSGFGRS
 LTETEANLDSYQTALEEVLWLLSIAEDLAQAQGEISNDVEEVEKQFHTHBGVMMDLTS
 HQGRVGNVLQGLSGLITGKLSDEDETEVQEQMNLNSRWELRVASMKQSNLHKVL
 MDLQNVQLKELNDWLTGTEERTKWEKEPLGPDIEDLKRQVQHKQLQEDLEQEQVRV
 NSLTHMVVVVDESSGDHATAALEBQLKVLGDRWANICRWTEDRWVLLQDILLKWQRF
 EEQCLFSAMLSEKEDAVNKIHTTGPKDQSEVLSNLQKLAVLKLTLEKKQKTMDKLCSL
 NQDLSALKNTRVIAHKMEAWLNFQARQWNLQVLEKSSAQISQAVTTQPSLTKTTV
 METVMTVTTREHILVKHAQEEPLPPPPQKKRQIIVDESEIRKELDQVITELHLSWITRSE
 ADVQSPFEFAYRKEGNFSDLKEKYNATIEREAKFKRLQDASRSQAALVQVMVNEGVN
 ADSIKQASEQNSRWIEFCQLLSERLNWLEYQNNIITFYNLQQLQLEQMTTAAENWLKT
 QPTTTFSEPTAKSQLKICKDEINRLSPMQYQIERLKIQISIALKEKGQPMFLDFFVTA
 FTFHFNQVFDVQAREKELQTIIFSLPALPMRYQETSMILTIIQQOSETKLSIPQVTVTE
 YDIMEQRLGADVLQSSLEQOQNGNLYLSTTVKEMSKAPLSISRKYSBFEIEEGR
 WKLSSQLVEHCQKLEEQMAKRLKIQNHKTLKKWITEVDVFLKEEWPALGDSLEILKR
 QLKOCRLLVNDQTIQPSLNSVNEGAQWKNEAPEEPAGLELETRELNTQWDMCRQ
 VYARKEALKGGLDKTVSLQKLSSEVHEWMTQAEELYLBERDFTYKTPDELQTAVEEMKR
 AKEEAQKEAKVKLLTQSVNSVIAQAPPAQAAKKKELDTLTNYQWCLRNLGCKKT
 LEEVWACWHELLSYLEKANKWLSEVVKLKTENISGAEBIEAEVLDLENLMQHSED
 NPNQIRILAQTLTDGGVMDELINEELTFNSRWRELHEEAVRRQKLEQSTQSAQIEIE
 KSLHLIQESLSSDKQLAAYIADKYDAAQMPQEAQIKQSDLTSHEISLEEMKHNQGGK
 ETAQRVLSQIDVAKQKLDVMSKFLQPKPAENQEQKLSKMWILDEVKMLPALETKS
 VEQEVQSOINHCNLYKLSSEKSEVEMVIKTRQIQVQKQTEENKELDERVTALIK

CDs

500

	Query Match	51.5%;	Score 1030.6;	DB 4;	Length 13887;
	Best Local Similarity	91.7%;	Pred. No. 1.7e-247;		
	Matches 1090;	Conservative	0;	Mismatches 99;	Indels 0; Gaps 0;
QY	813	GGTACCTACTCATAGATTACTGCAACAGATTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG	872		
Db	8326	GGAAAGAAACCCACAGATTACTGCAACAGATTCCCCCTTGGACCTGGAGAAGTTCCCTTGCCTG	8385		
QY	873	GCTTACAGAAAGTGAAAAACAACCTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT	932		
Db	8386	GCTTACAGAAAGCCGAAAAACAACCTGCCAACGTCTCTGCAGGATGCCACCCATAAGGAAAGGCT	8445		
QY	933	CCTAGAAGACTCCAAAGGGAGTAAAGAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGTGA	992		
Db	8446	TCTAGAAGATTCCAAAGGGAGTAAGAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGGAGA	8505		
QY	993	AATTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG	1052		
Db	8506	AATCGAAGCTCACACAGATATCTATCAACAACCTGGACGAAAAATGGCCAAAAAGTCCTGAG	8565		
QY	1053	ATCCCTGGAAGGTTCCGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTT	1112		
Db	8566	ATCCCTGGAAGGTTCTGACGATGCAGCCTTGTGTGCAAGACGTTTGGATAACATGAACCTT	8625		
QY	1113	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTC	1172		
Db	8626	CAAGTGGAGCGAACTTCGGAAAAAGTCTCTCAACATTAGGTCTCACTTGGAAAGCCAGTTC	8685		
QY	1173	TGACCAGTGGAAAGGCTCTGCACCTTTCTCTCTGCAGGAACCTCTGGTGTGGCTACAGCTGAA	1232		

b 8686 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTTCAGGAACCTTCTGGTATGGCTCCAGCTGAA 8745
y 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTATTGGAGCGGACCTTTCCAGCAGTTTCAAGAAGCA 1292
b 8746 AGATGATGAGTTAAGCCGGCAGGCACCCATTGGAGGAGACTTTCAGCGGTGCAAGCA 8805
y 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATATCATGAG 1352
b 8806 GAATGATGTACACAGGGCCCTTCAAGAGGGAATTGAAAACGAAAGAACCTGTATATCATGAG 8865
y 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
b 8866 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTAGAAGGACTAGAGAAACT 8925
y 1413 CTACAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
b 8926 CTACAGGAGCCAGAGAGCTGCCCTCTGAAAGAGAGAGAGCCAGAAATGTCACTCGGCTCCT 8985
y 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCGGCTGA 1532
b 8986 ACGAAAGCAAGCTGAGGAGGTCAACACTGAGTGGGAAAAAATTGAACGTGCACTCTGCAGA 9045
y 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
b 9046 CTGGCAGAGAAAAATAGACGAGGCCCTCGAAAGACTCCAGAGCTTCAGGAAGCAACAGA 9105
y 1593 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGG 1652
b 9106 TGAGCTGGATCTCAAACTACGTGAGGAGAGGTGATCAAAAGGATCCTGGCAGCCCTGTGGG 9165
y 1653 CGATCTCCTCAATGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
b 9166 TGACCTCCTCAATGACTCTCTCCAAAGATCACCTCGAAAGAGTCAAGGCCTTCGAGGAGA 9225
y 1713 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
b 9226 AATTACACCTCTGAAAGAGAAATGTGACGTACGTCAATGACCTTGCTCGCCAACTCACTAC 9285
y 1773 TTTGGGCATTGAGCTCTCACCCTGATTAACCTCAGCACCTCTGGAAGACCTGAAACACAGATG 1832
b 9286 GTTGGGCATTGAGCTGTACCATATTAACCTCAACACTCTGGAAGACCTGAACACAGATG 9345
y 1833 GAAGCTTCTGCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGA 1892
b 9346 GAAGCTTCTGCAGTGGCCATTGAGGACCGCATCAGGCAGCTGCAATGAAGCCACAGGGA 9405
y 1893 CTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCAGCTCTGTCCAGGTCCTCGGGAGAGAGC 1952
b 9406 CTTTGGACCAAGCTCCAGCAGCTTCTTTCCAGCTTCTGTCCAGGTCCTCGGGAGAGAGC 9465
y 1953 CATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCAGAGACTCAAAACA 2001
b 9466 CATCTCACCACCAACAAAGTGGCCCTACTATATCAACCCAGAGACCCAAACA 9514

ESULT 14
X306153
OCUS
EFINITION
Sequence 904 from Patent WO0188188.
CESSION
AX306153
ERSION
AX306153.1 GI:17645441
EYWORDS
Mus musculus (house mouse)
OURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE
Method for examining ischemic conditions
JOURNAL
Patent: WO 0188188-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1. .13815
source

/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 50.2%; Score 1004; DB 6; Length 13815;
Best Local Similarity 90.3%; Pred.No. 8.2e-241;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
Db 8251 GGAAGAAACTCATAGATTACTGACGAGTTCCCTCTGGACCTGGAGAAAGTTTCTTTCCTG 8310
QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGGCT 932
Db 8311 GATTACGGAAGCAGAAAAAACTGCCAATGTCTCTACAGGACGCTTCCCGTAAGGAGAAGCT 8370
QY 933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA 992
Db 8371 CCTAGAAGACTCCAGGGAGTCAGAGAGCTGATGAAACCAATGGCAAGATCTCCAAGGAGA 8430
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAG 1052
Db 8431 AATTGAAGCTCACACAGATATCTATCACAACTGTGATGAAAAATGGCCAAAAAATCCTGAG 8490
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCAATTGGAAAGCCAGTTT 1112
Db 8491 ATCCCTGGAAGGTTCCGATGATGAGCAGCCCTCTGTAACAAAGACGTTTGGATAAATCATGA 8550
QY 1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCAATTGGAAAGCCAGTTT 1172
Db 8551 CAAGTGGAGTGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCCAATTGGAAAGCAAGTTT 8610
QY 1173 TGACAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTTGGTGTGCTACAGCTGAA 1232
Db 8611 TGACAGTGGAAAGCGTCTGCATCTTTCTCTCAGGAACTTCTTGTGGTGTGCTACAGCTGAA 8670
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTTCGAGGCGGACTTTCAGCAGTTTCAAGAGCA 1292
Db 8671 AGATGATGAATTAAGCCGGTTCAGGACCCCTCGTGGTGTGATTTCCAGCAGTTTCAAGAGCA 8730
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATGAAAACTAAAGAACCTGTAATCATGAG 1352
Db 8731 GAATGATGTACATAGGGCCCTTCAAGAGGGAATGAAAACTAAAGAACCTGTAATCATGAG 8790
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACT 1412
Db 8791 TACTCTTGAGACTGTGAGAAATATTTCTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACT 8850
QY 1413 CTACAGGAGCCAGAGAGCTGCCCTCTCTGAGGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 8851 CTACAGGAGCCAGAGAACTGCCCTCTCTGAAAGAAAGAGCTCAGAAATGTCACTCGGCTCCT 8910
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 1532
Db 8911 ACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGACAAATTGAACCTGCGCTCAGCTGA 8970
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
Db 8971 TTGGCAGAGAAAAATAGATGAAGCTCTTGAAGACTCCAGGAACTTCAAGAGACTGCCGA 9030
QY 1593 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCGCTGGG 1652
Db 9031 TGAACCTGGACCTCAAGTTTCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCAGTGGG 9090
QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGA 1712
Db 9091 GGATCTCCTCATTTGACTCTCTGCAAGATCACCTTGAAAAAGTCAAGGCACTTCGGGGAGA 9150
QY 1713 AATTGCGCCTCTGAAAGAGAAAGTGAAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCAC 1772
Db 9151 AATTGCACTCTTAAAGAGAAATGTCAATCGTGTCAATGACCTTGCATCATGAGCTGACCAC 9210

1773 TTTGGGCAATTCAGCTCTCACCGTATAAACCTCAGCACTCTGTGAAGACCTTGAACACCAGATG 1832
9211 ACTGGGCAATTCAGCTCTCACCTTATAAACCTCAGCACTTTGGAAGATCTGAATACCAGATG 9270
1833 GAAGCTTCTCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAGGGA 1892
9271 GAGGCTTCTACAGGTGGCTGTGGAGGACCGTGTGAGAGCTGCATGAAGCCACAGGGA 9330
1893 CTTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTCTGCCAGGTCCTCGGAGAGAGC 1952
9331 CTTTGGTCTGCATCCAGCACTTCTTTCCACTTCAGTTCAGGTCCTCGGAGAGAGC 9390
1953 CATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAC 2000
9391 CATCTCACCNAACAAAGTGCCTTACTATATCAACCACGAGACCCAAAC 9438

RESULT 15
AX538582
LOCUS AX538582 13815 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX538582
VERSION AX538582.1 GI:25271088
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Chamberlain J.S. and Harper S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 2 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..13815
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 50.2%; Score 1004; DB 6; Length 13815;
Best Local Similarity 90.3%; Pred. No. 8.2e-241;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCCTG 872
8251 GGAAGAACTCATAGATTACTGCAGCAGTTCCCTCTGGACCTGGAGAGTTTCTTCTCTG 8310
873 GCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTGAAGAAAGGCT 932
8311 GATTACGGAAGCAGAAACAACTGCCAATGTCTTACAGGACGCTTCCCGTAAGGAGAAGCT 8370
933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
8371 CCTAGAAGACTCCAGGGGAGTCAGAGAGCTGATGAAACCATGGCAAGATCTCCAAGGAGA 8430
993 AATTGAAGCTCACACAGATGTTTATCAAAACCTGGATGAAACAGCCAAAATCCTGAG 1052
8431 AATTGAAACTCACACAGATATCTATCACAATCTTGATGAAAATGGCCAAAATCCTGAG 8490
1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTTACAAAGACGTTTGGATAACATGAACCTT 1112
8491 ATCCCTGGAAGGTTCCGATGATGAAGCACCCTGTTTACAAAGACGTTTGGATAACATGAATT 8550
1113 CAAGTGGAGTGAACCTTCGAAAAGTCTCTCAACATTAGTCCCATTTTGGAGGCAGTTTC 1172
8551 CAAGTGGAGTGAACCTTCAGAAAAGTCTCTCAACATTAGTCCCATTTTGGAGGCAGTTTC 8610
1173 TGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
8611 TGACCAGTGAAGCGTTTGCATCTTTTCTCTTCAGGAACCTTCTGTTTGGCTACAGCTGAA 8670
1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCCAGCAGTTTCAGAAGCA 1292

Search completed: April 5, 2004, 02:39:39
Job time : 5220.76 secs

8671 AGATGATGAATGAGCCGTGAGGCAACCCATCGGTGGTGATTTCCAGCAGTTTCAGAAGCA 8730
1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTATCATGAG 1352
8731 GAATGATATACATAGGGCCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTATCATGAG 8790
1353 TACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
8791 TACTCTGGAGACTGTGAGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 8850
1413 CTACCAGGAGCCACAGAGAGCTGCCTCTCTGAGGAGAGAGCCACAGAAATGTCACTCGGCTTCT 1472
8851 CTACCAGGAGCCACAGAGAACTGCCTCTCTGAAGAAAGAGCTCAGAATGTCACTCGGCTCCT 8910
1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA 1532
8911 ACGAAAGCAGGCTGAAGAGGTCAACCGTGAATGGGACAAATTGAACCTGCGCTCAGCTGA 8970
1533 CTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
8971 TTGGCAGAGAAAAATAGATGAAGCTCTTGAAGAACTCCAGGAACTTCAGGAAGCTGCCGA 9030
1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
9031 TGAACCTGACCTCAAGTTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCAGTGGG 9090
1653 CGATCTCCTATTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
9091 GGATCTCCTCATGACTCTCTGCAAGATCACCTTGAAAAAGTCAAGGCACCTTCGGGGAGA 9150
1713 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCAATGACCTTGCTGCTGCCAGCTTACCAC 1772
9151 AATTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAATGACCTTGACATCAGCTGACCAC 9210
1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
9211 ACTGGGCATTGAGCTCTCACCTTTATAACCTCAGCACTTTTGAAGATCTGAATACCAGATG 9270
1833 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 1892
9271 GAGGCTTCTACAGGTGGCTGTGGAGGACCGTGTGACACAGCTGCATGAAGCCCACAGGGA 9330
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
9331 CTTTGGTCTGCATCCAGCACTTCTCTTCCACTTCCAGTTTCCAGGGTCCCTGGGAGAGAGC 9390
1953 CATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAC 2000
9391 CATCTCACCNAACAAAGTGCCTTACTATATCAACCACGAGACCCAAAC 9438

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

4 nucleic - nucleic search, using sw model
on on: April 4, 2004, 16:52:43 ; Search time 3673.37 Seconds
(without alignments)
17079.796 Million cell updates/sec

File: US-09-845-416-2_COPY_900_3000
Effect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcactcttgaagac 2101

oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
arched: 27513289 seqs, 14931090276 residues
otal number of hits satisfying chosen parameters: 55026578

imum DB seq length: 0
imum DB seq length: 2000000000
st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

itabase : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1103.4	52.5	3870	11	BC036103 Homo sapi
2	866.6	41.2	3056	11	AK044536 Mus muscu
3	846.2	40.3	5691	29	AY399453 Homo sapi
4	633.4	30.1	5697	29	AY399455 Mus muscu

5	597.2	28.4	728	14	CB228986	CB228986	AGENCOURT
6	592.4	28.2	834	12	BI729851	BI729851	603349511
7	514.4	24.5	5676	29	AY399454	AY399454	Pan trogl
c 8	509	24.2	595	14	CB177816	CB177816	is21c01.x
9	493.2	23.5	1047	14	CB850319	CB850319	MRA-0070
10	491.6	23.4	750	12	BI730168	BI730168	603349711
11	486.6	23.2	579	9	AL121550	AL121550	DKEZP762L
12	424.2	20.2	663	12	BM488464	BM488464	pgm2n.pk0
13	389	18.5	644	13	BU313510	BU313510	603540290
14	343	16.3	2334	11	BC011062	BC011062	Mus muscu
15	334.6	15.9	9915	29	AY407022	AY407022	Homo sapi
c 16	324.6	15.4	402	14	CB547284	CB547284	AMGNNUC:S
17	312.2	14.9	9691	29	AY407024	AY407024	Mus muscu
18	311.8	14.8	9096	29	AY407023	AY407023	Pan trogl
19	275.2	13.1	1122	14	CF109978	CF109978	Shultzomi
20	265.8	12.7	772	12	BI250598	BI250598	602993659
21	226	10.8	250	10	BF963618	BF963618	QV2-NN004
22	217.2	10.3	851	13	BU201022	BU201022	603952191
23	213.2	10.1	2874	29	AY408546	AY408546	Homo sapi
24	212.2	10.1	3051	11	BC036095	BC036095	Homo sapi
25	203.6	9.7	1541	11	AK034383	AK034383	Mus muscu
26	203.6	9.7	2874	29	AY408548	AY408548	Mus muscu
27	203.6	9.7	3753	11	AK081426	AK081426	Mus muscu
28	192	9.1	778	14	CB524596	CB524596	UI-M-FY0-
29	184.4	8.8	812	14	CF748008	CF748008	UI-M-HE0-
30	183	8.7	763	14	CB518960	CB518960	UI-M-GH0-
c 31	182.4	8.7	645	10	AW467977	AW467977	he29g08.x
c 32	177	8.4	646	12	BI289102	BI289102	UI-R-DK0-
33	172.8	8.2	835	12	BI553820	BI553820	603190772
34	172.8	8.2	2874	29	AY408547	AY408547	Pan trogl
35	166.2	7.9	907	13	BU122401	BU122401	603003073
36	165.2	7.9	600	12	BI988528	BI988528	4012-24 M
c 37	162	7.7	427	10	BE817979	BE817979	CM2-BN027
c 38	158.2	7.5	569	29	CE228097	CE228097	tigr-gss-
39	158.2	7.5	615	14	CA377239	CA377239	655768 NC
40	153.2	7.3	575	29	CG519934	CG519934	OST84444
41	153.2	7.3	885	11	AK020881	AK020881	Mus muscu
c 42	150.6	7.2	515	28	AZ780914	AZ780914	2M0018O10
43	150.2	7.1	497	29	CG525393	CG525393	OST100373
44	150.2	7.1	497	29	CG653669	CG653669	OST419391
45	149.2	7.1	497	29	CG553287	CG553287	OST165637

ALIGNMENTS

RESULT 1
BC036103 3870 bp mRNA linear HTC 19-NOV-2003
LOCUS
DEFINITION
Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (cdna clone IMAGE:5274415), with apparent retained intron.
ACCESSION
BC036103
VERSION
BC036103.1 GI:23271310
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3870)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalrus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
12477932
PUBMED 2 (bases 1 to 3870)
REFERENCE Strausberg,R.
AUTHORS Direct Submission
TITLE
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: retained intron.
FEATURES
source
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:527415"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

Query Match 52.5%; Score 1103.4; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 8.2e-235;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGGAGCCCAAT 60
1043 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGGAGCCCAAT 1102
61 TCCTTCACAGCATTGGAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 120
1103 TCCTTCACAGCATTGGAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 1162
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
1163 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1222
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGAAAGTGGTGAAGA 240
1223 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGAAAGTGGTGAAGA 1282
241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 300
1283 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1342

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1343 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1402
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1403 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1462
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 1463 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAAATCAGAAACT 1522
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 1523 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 1582
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
Db 1583 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1642
QY 601 AGAAGATCTAGAACACAAGCAAGTCAAGGTCAATCTCTCACTCAGATGGTGGTGGTAGT 660
Db 1643 AGAAGATCTAGAACACAAGCAAGTCAAGGTCAATCTCTCACTCAGATGGTGGTGGTAGT 1702
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTTGAAGAAACAACCTAAGGTATTGGG 720
Db 1703 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTTGAAGAAACAACCTAAGGTATTGGG 1762
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 780
Db 1763 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1822
QY 781 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTGTAGTGGCTTTTCAGA 840
Db 1823 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTGTAGTGGCTTTTCAGA 1882
QY 841 AAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAATGTT 900
Db 1883 AAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAATGTT 1942
QY 901 ATCAAGTCTTCAAAAACCTGGCGTTTAAAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
Db 1943 ATCAAGTCTTCAAAAACCTGGCGTTTAAAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 2002
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACACTGAAGATAAAGTCAGTGAC 1020
Db 2003 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACACTGAAGATAAAGTCAGTGAC 2062
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 2063 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAAA 2122
QY 1081 ACTTGAAGAGATACAGCACAGACT 1105
Db 2123 ACTTGAAGAGATACAGCACAGATT 2147

RESULT 2
AK044536
LOCUS
DEFINITION
AK044536 3056 bp mRNA linear HTC 20-SEP-2003
Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930019F21 product:dystrophin, muscular dystrophy, full
insert sequence.
ACCESSION AK044536
VERSION AK044536.1 GI:26090404
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.

b 1784 AAAAGAAGATGCAATGAAGAACAATTCAGACAAAGTGGCTTTAAAGATCAAAAATGAATGAT 1843
Y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
b 1844 GTCAAGTCTTCAAAAATATCTACTTTAAAAATAGATCTAGAAAAGAAAAAGCAACCAT 1903
Y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
b 1904 GGAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAAATAAGTCAGTGAC 1963
Y 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
b 1964 TCAAAAGATGGAATCTGGATGGAAAACTTTGCACAACGTTGGGACAATTTAACCAAAA 2023
Y 1081 ACTTGAAGAAGATACAGACAGACT 1105
b 2024 ACTTGAAGAAGATTCAGACAAAATT 2048

ESULT 3
Y399453 AY399453 5691 bp DNA linear GSS 12-DEC-2003
OCUS Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
EFINITION genomic survey sequence.
CCENSION AY399453
ERSION AY399453.1 GI:39755442
EYWORDS GSS.
OURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5691)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
EFERENCE 2 (bases 1 to 5691)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
OMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
EATURES Location/Qualifiers
source 1..5691
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..5691
/locus_tag="HCM0229"

RIGIN
Query Match 40.3%; Score 846.2; DB 29; Length 5691;
Best Local Similarity 89.1%; Pred. No. 1.7e-177;
Matches 985; Conservative 0; Mismatches 3; Indels 117; Gaps 2;

Y 1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCAT 60
b 294 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCAT 353

Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
b 354 TCCTTCACAGC-----TCCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 404
Y 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATATCGTGGCTTCTTTC 180

RESULT 4
AY399455
LOCUS AY399455 5697 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

Db 405 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 464
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 465 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 524
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 525 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 584
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGGAAAAATTATCAGAAGATGAAGA 360
Db 585 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGGAAAAATTATCAGAAGATGAAGA 644
QY 361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 645 AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 704
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAAC 480
Db 705 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAAC 764
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACACAGGAAAAATGGAGGAAGA 540
Db 765 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACACAGGAAAAATGGAGGAAGA 824
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAACAAACATAAGGTGCTTCA 600
Db 825 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAACAAACATAAGGTGCTTCA 884
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCGAATTTCTCTCACTCACATGGTGGTGTAGT 660
Db 885 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCGAATTTCTCTCACTCACATGGTGGTGTAGT 944
QY 661 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 945 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1004
QY 721 AGATCGATGGGCCAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
Db 1005 AGATCGATGGGCCAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 1064
QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACACAGTGCCTTTTTTACTGCATGGCTTTTCA 840
Db 1065 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAG----- 1098
QY 841 AAAAGAAGATGCAGTCAACAAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1099 ----- 1098
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 1099 -----GTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1136
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACACTGAAGAATAAGTCAGTGAC 1020
Db 1137 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACACTGAAGAATAAGTCAGTGAC 1196
QY 1021 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 1197 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1256
QY 1081 ACTTGAAGAAGATACAGCACAGACT 1105
Db 1257 ACTTGAAGAAGATACAGCACAGATT 1281

CESSION AY399455
ERSION AY399455.1 GI:39755444
YWORDS GSS.
URCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5697)
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..5697
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>5697
/locus_tag="HCM0229"
Gene
IGIN
Query Match 30.1%; Score 633.4; DB 29; Length 5697;
Best Local Similarity 77.1%; Pred. No. 4e-130;
Matches 852; Conservative 0; Mismatches 136; Indels 117; Gaps 2;
1 GAGCTATGCTACACACAGGCTGCTATGTCAACCACCTCTGACCCCTACACGAGCCCAT 60
300 GAGTTATGCTTACACAGGCTGCTATGTTGCCACCTCTGATTCACACAGAGCCCTA 359
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATGATGGAGAG 120
360 TCCTTCACAG-----GCTCCAGAGACAAGTCACTTGACAGTTCATGATGGAGAC 410
121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTCTTTC 180
411 GGAAGTAAATCTGGATAGTTACCAAACCTGCTTTAGAGAAGTACTTTTCATGGCTCTTTC 470
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
471 TGCCGAGGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 530
241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTACAGCCCATCAGGCCGGGTTGG 300
531 ACAGTTTCATGCTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 590
301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
591 TAATGTTCTACAGTTAGGAAGTCAAACTAGTTGGAAAAGGGAAATTTATCAGAAGATGAAGA 650
361 AACTGAATPACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 420
651 AGCTGAATGCAAGAACAATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 710
421 TAGCATGGAACAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAATCAGAAACT 480
711 TAGCATGGAACAAACAAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAATCAGAAAT 770
481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA 540

Db 771 AAAAGAACTAGATGACTGGTTAAACAAAAAAGTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 830
Qy 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACAGCCCAAGTACAAACAACATAAGGTGCTTCA 600
Db 831 GCCCTTTGGACCTGATCTTTGAAGATCTAAATGCCAAAGTACAAACAACATAAGGTGCTTCA 890
Qy 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAGGTCATCTCTCACTCACATGGTGGTGTAGT 660
Db 891 AGAAGATCTAGAACAGGAGCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAG 950
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 951 TGATGAATCCAGCGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 1010
Qy 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTCTTTTACAAGACAT 780
Db 1011 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTTTACAAGATAT 1070
Qy 781 CCTTCTCAAATGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 840
Db 1071 TCTTCTAAATGCGAGCATTTTACTGAGAACAG----- 1104
Qy 841 AAAAGAAAGATGCAAGTGAACAAGATTCAACAACAGTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1105 ----- 1104
Qy 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
Db 1105 -----ACTTTAAAAATAGATCTAGAAAAAGAAAAGCCCAACCAT 1142
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 1143 GGAAAAAATAAGTTCATCAATCAAGATCTACTTTGGGCACTGAAAAATAAGTCAGTGAC 1202
Qy 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGCTTGGGATAATTTAGTCCAAAA 1080
Db 1203 TCAAAAGATGGAAATCTGGATGGAAACTTTTGCAACACGTTGGGACAATTTAACCCAAAA 1262
Qy 1081 ACTTGAAAAAGATGACAGCACAGACT 1105
Db 1263 ACTTGAAAAAGATTGAGCACAAAT 1287
RESULT 5
CB228986
LOCUS
DEFINITION CB228986 728 bp mRNA linear EST 10-FEB-2003
AGENCOURT 11499247 NICHD Rh.Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.
1 (bases 1 to 728)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM3135 row: e column: 11
High quality sequence stop: 583.
FEATURES
Location/Qualifiers

source 1. .728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHHD Rh Ovi"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHHD Library."

ORIGIN

Query Match 28.4%; Score 597.2; DB 14; Length 728;
Best Local Similarity 97.0%; Pred. No. 3.1e-122;
Matches 619; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

2Y 295 GGTGGTAATATTCTACAAATGGGAAGTAAGCTGATTGGACAGGAAATATATCAGAAGA 354
|||||
Db 1 GGTGGTAATATTCTACAAATGGGAAGTCAGCTGATTGGACAGGAAATATATCAGAAGA 60
|||||

2Y 355 TGAAGAACTGAAGTACAAAGACAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAG 414
|||||
Db 61 TGAAGAACTGAAGTACAAAGACAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAG 120
|||||

2Y 415 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA 474
|||||
Db 121 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA 180
|||||

2Y 475 GAAACTGAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGCAAGGAAATGGA 534
|||||
Db 181 GAAACTGAAGAGTTGAATGACTGGCTGACAAAACAGAAAGAAAGCAAGGAAATGGA 240
|||||

2Y 535 GGAAGACCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACACAAATAGGT 594
|||||
Db 241 GAAAGACCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACACAAATAGGT 300
|||||

2Y 595 GCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGT 654
|||||
Db 301 GCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGT 360
|||||

2Y 655 GGTAGTTGATGAATCTAGTGAGATCAGGCAACTGCTGCTTTGGAGAACTTAAGGT 714
|||||
Db 361 GGTAGTTGATGAATCTAGTGAGATCAGGCAACTGCTGCTTTGGAGAACTTAAGGT 420
|||||

2Y 715 ATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACA 774
|||||
Db 421 ATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACA 480
|||||

2Y 775 AGACATCCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCCTTTTAGTGATGGCT 834
|||||
Db 481 AGACATCCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCCTTTTAGTGATGGCT 540
|||||

2Y 835 TTCAGAAAAAGAGATGAGTGAACAAAGATTCACAACT-GGCTTTAAAGATCAAAATG 893
|||||
Db 541 TTCAGAAAAAGAGATGAGTGAACAGGATTCACAACTGGGCTTTAAGGATCAAAATG 600
|||||

2Y 894 AAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAAA 931
|||||
Db 601 AAATGTTATCAAGTCTTCAAAAACCTGGCCCGTATGTAA 638
|||||

RESULT 6
3I729851
LOCUS BI729851 834 bp mRNA linear EST 20-SEP-2001
DEFINITION 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
mRNA sequence.
ACCESSION BI729851
VERSION BI729851.1 GI:15706864
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11908 row: e column: 03
High quality sequence stop: 796.
Location/Qualifiers
1. .834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357162"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 28.2%; Score 592.4; DB 12; Length 834;
Best Local Similarity 88.4%; Pred. No. 3.7e-121;
Matches 677; Conservative 0; Mismatches 86; Indels 3; Gaps 3;

QY 1065 ATAATTTAGTCCAAAACTTGAAGAAGATACAGCAGACTCATAGATTACTGCAACAGT 1124
|||||
Db 44 ATAAAAGAGTAAGTGAGCAAGAGGCTGCTTTGGAGAAACTCATAGATTACTGCAGCAGT 103
|||||

QY 1125 TCCCCCTGGACCTGGAAAAAGTTTCTTGCCTGGTTACAGAAAGCTGAAACAACTGCCAATG 1184
|||||
Db 104 TCCCTCTGGACCTGGAGAAAGTTCTTCTCTGGATTACGGAAGCAGAAACAACTGCCAATG 163
|||||

QY 1185 TCCTACAGATGCTACCCGTAAAGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGC 1244
|||||
Db 164 TCCTACAGACGCTTCCCGTAGGAGAAAGCTCCTAGAAGACTCC-AGGAGTCAAGAGC 222
|||||

QY 1245 TGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACA 1304
|||||
Db 223 TGATGAAACCATGGCAAGATCTCCAAGGAGAAATTGAAGACTCACACAGATATCATCACA 282
|||||

QY 1305 ACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCC 1364
|||||
Db 283 ATCTTGATGAAAAATGGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGAAGCACCCC 342
|||||

QY 1365 TGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTC 1424
|||||
Db 343 TGTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACCTTCAGAAAAAGTCTC 402
|||||

QY 1425 TCAACATTAGGTCCCATTTTGGAAAGCCAGTTTCTGACCAGTGAAGCGTCTGCACCTTTCTC 1484
|||||
Db 403 TCAACATTAGGTCCCATTTTGGAAAGCAAGTTCTTGACCAGTGAAGCGTTTGCATCTTTCTC 462
|||||

QY 1485 TGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTA 1544
|||||
Db 463 TTCAGGAACCTTCTTGTGGTGGCTACAGCTGAAAGATGATGAACCTGAGCCGTCAGGCACCCA 522
|||||

QY 1545 TTGGAGGCGACTTTCAGCAGCTTCAGAAAGCAGAACCATGTACATAGGGCCCTTCAAGAGGG 1604
|||||
Db 523 TCGGTGGTGATTTCCAGCAGTTCAGAAAGCAGAAATGATATACATAGGGCCCTTCAAGAGGG 582
|||||

Y 1605 AATTGAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGA 1664
b 583 AATTGAAACTAAAGAACCTGTATCATGAGTACTCTGGAGACTGTGAGATATTCTGA 642
Y 1655 CAGAGCAGCCTTTGGGAAGGACTAGAGAACTC-TACCAGGAGCCCGAGAGCTGCCTCCT 1723
b 643 CAGAGCAGCCTTTGGGAAGGACTAGAGAACTCTTACCAGGAGCCCGAGAGCTGCCTCCT 702
Y 1724 GAGGAGAGAGCCGAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1783
b 703 GAAGAAAGAGCTCAGAAATGTCACTCGGCTCCTACGAAAGCAGGCTGAAGAGGTCAACGCT 762
Y 1784 GAGTGGGAAATTTGAACTGCACTC-CGCTGACTGGCAGAGAAA 1828
b 763 GAATGGGACCAATTGAACTCGGCTCAAGCTGATTGGCAGAGAAA 808
RESULT 7
Y399454
OCUS
EFINITION Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
CCESION AY399454
ERSION AY399454.1 GI:39755443
EYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 5676)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5676)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..5676
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5676
/locus_tag="HCM0229"
RIGIN
Query Match 24.5%; Score 514.4; DB 29; Length 5676;
Best Local Similarity 75.2%; Pred. No. 1.3e-103;
Matches 636; Conservative 0; Mismatches 102; Indels 108; Gaps 1;
Y 260 GGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGGTAATATTCTACAATTGGGA 319
b 529 GGGTACATGATGATNNNNAGCCCATCAGGCGCGGTGGTAATANNNNNNANTTGGGA 588
Y 320 AGTAAGCTGATTGGAAACAGGAAAATATTCAGAAGATGAAGAACTGAAGTACAAGAGCAG 379
b 589 AGTNNNTGATTGGAAACAGGAAAATATTCAGAAGATGAAGAACTGAAGTACAAGAGCAG 648
Y 380 ATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGTAGCATGGAACAAAGC 439
b 649 ATGAATCTNNNNNATTCAGATGGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 708

QY 440 AATTTCATAGAGTTTAAATGGATCTCCAGAACTCAGAAACTGAAAGAGTTGAATGACTGG 499
Db 709 NNTTTCATAGAGTTTAAATGGATCTCCAGAACTCAGAAACTGAAAGAGTTGAATGACTGG 768
QY 500 CTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAGAGCCCTCTTGGACCTGATCTT 559
Db 769 CTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAGAGCCCTCTTGGACCTGATCTT 828
QY 560 GAAGACCTAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAGAA 619
Db 829 GAAGACCTAAACGCCAAGTACAACAACATAAGNNNNNNNCAAGAAAGATTTAGAACAAGAA 888
QY 620 CAAGTCAGGTCATTTCTCTCACTCACATGCTGGTGGTAGTTGATGAATCTAGTGGAGAT 679
Db 889 CAAGTCAGGTCATTTCTCTCACTCACATGCTGGTGGTAGTTGATGAANNNNANTGGAGAC 948
QY 680 CACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATC 739
Db 949 CACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATC 1008
QY 740 TGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGT 799
Db 1009 TGTAGTGGACAGAAAGCCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGT 1068
QY 800 CTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGAAAAAGAGATGCAGTGAAC 859
Db 1069 CTTACTGAAGAACAG----- 1083
QY 860 AAGATTCACACAACACTGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTG 919
Db 1084 ----- 1083
QY 920 GCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTC 979
Db 1084 ---GTTTAAAGCGGATNNNGAAAGAAAAGCAATCCATGGGCAAACTATATTNNNN 1140
QY 980 AAACAAGATCTTCTTCAACACACTGAAGATAAGTCACTGACCCAGAACGCGAAGCATGG 1039
Db 1141 NNACAAGTNTTCTNTNAACACTGANGAATAAGTCACTGACCCNNNAAGACGGAAGCNGG 1200
QY 1040 CTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAACCTGAAAAGAGTACAGCA 1099
Db 1201 NNGGATAACTTTGCCCGNNNTTGGGATAATTNNNCCAAAACNNGAAAAGAGTACAGCA 1260
QY 1100 CAGACT 1105
Db 1261 CAGATT 1266
RESULT 8
CB177816 595 bp mRNA linear EST 31-JAN-2003
LOCUS is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
DEFINITION similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
ACCESSION CB177816
VERSION CB177816.1 GI:28186206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lienishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohph.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
Location/Qualifiers
1. 595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

FEATURES

source

ORIGIN

Query Match 24.2%; Score 509; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.3e-102;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCACACACCTCTGACCCCTACACGGAGCCCAT 60
|||||
509 GAGCTATGCCTACACACAGGCTGCTTATGTCACACACCTCTGACCCCTACACGGAGCCCAT 450
|||||
61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAGCAAGTCATTTGGCAGTTCATTGATGGAGAG 120
|||||
449 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAGCAAGTCATTTGGCAGTTCATTGATGGAGAG 390
|||||
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 180
|||||
389 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 330
|||||
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
329 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 270
|||||
241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
|||||
269 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 210
|||||
301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
|||||
209 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 150
|||||
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
|||||
149 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 90
|||||
421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
89 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 30
|||||
481 GAAAGAGTTGAATGACTGGCTAAACAAAA 509
|||||
29 GAAAGAGTTGAATGACTGGCTAAACAAAA 1
|||||

RESULT 9

CB850319
LOCUS
DEFINITION MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION CB850319
VERSION CB850319.1 GI:34380806
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1047)
Yu, J., Farjo, R., MacNee, S.P., Baehr, W., Stambolian, D.E. and
Swaroop, A.
TITLE Annotation and analysis of 10,000 expressed sequence tags from
developing mouse eye and adult retina
JOURNAL Genome Biol. 4 (10), R65 (2003)
MEDLINE 22881944
PUBMED 14519200
COMMENT Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu.
Location/Qualifiers
1. 1047
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="retina"
/clone_lib="MOUSE ADULT RETINA"
/note="Vector: pSPORT1"

ORIGIN

Query Match 23.5%; Score 493.2; DB 14; Length 1047;
Best Local Similarity 83.2%; Pred. No. 4.8e-99;
Matches 610; Conservative 0; Mismatches 118; Indels 5; Gaps 5;
133 GGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACAC 192
|||||
86 GGATAGTTACCAAACTGCTTTANAAGAAGTACTTTTCATGGCTTCTTTCTGCGGAGGATAC 145
|||||
193 ATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGACCAAGTTTCATAC 252
|||||
146 ATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGACAGTTTCATGC 205
|||||
253 TCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGGTAAATATTCTACA 312
|||||
206 TCATGAGGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGTTGGTAAATGTTCTACA 265
|||||
313 ATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACA 372
|||||
266 GTTAGGAAGTCAACTAGTTGGAAAAGGGGAAATTTATCAGAAGATGAAGAGCTGAAGTGCA 325
|||||
373 AGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAG-CTAGCATGGAAA 431
|||||
326 AGAACAAATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTANCCTAGCATGGAAA 385
|||||
432 AACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGA 491
|||||
386 AACAAAGCAAAATTACACAAAGTTCTAATGGATCTCCAGAATCAGAAATTAAGAAGAACTAG 445
|||||
492 ATGACTGGCTAAACAAAAACAGAGAAAAGAACAAAGAAAATGGAGGAAGAGCCTCTTTGGAC 551
|||||
446 ATGACTGGTTAAACAAAAACTGAAGAGAGAACTAANAANAATGGAGGAAGAGCCTTTTGGAC 505
|||||
552 CTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGC-TTCAAGAAGATCTA 610
|||||
506 CTGATCTTGAAGATCTAAATATGCCCAAGTACAAACAATAAGGTGCTTTTCAAGAAGATCTA 565
|||||

611 GAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCT 670
|||||
566 GAACAGGAGCAGGTTCAGGGTCAACTCGCTCACTCACATGAGTAGTGGGTGATGAATCC 625
|||||
671 AGTGGAGATCAGCAACTGCTGCTTTTGGGAAGAACAACTTAAAGGTATTGGGAGATCGATG 729
|||||
626 AGCGGGGATCATGCAACAGCTGCTTTGGGAANAACAATTTAAGNACTGGGAGATCGATG 685
|||||
730 GGCAAAACATCTG-TAGATGGACAGAACAGCGCTGGGT-CTTTTACAAGACATCCTTCTC 787
|||||
686 GGCAATATCTGCAAAATGGACTGAAACCGTTGGATGTTTTTACANGATTTTCTTCTA 745
|||||
788 AAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTACAGAAAAGAA 847
|||||
746 AAATGGCAGCNTTTTACTGAAAAAANANNGCCTTTTATNCANGNTTTTAAAAAATAAAA 805
|||||
848 GATGCAGTGAACA 860
|||||
806 TGCATTGANAACA 818
|||||

BI730168 750 bp mRNA linear EST 20-SEP-2001
603349711P1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
mRNA sequence.

BI730168
BI730168.1 GI:15707181
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 750)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11908 row: f column: 04
High quality sequence stop: 747.
Location/Qualifiers
1. 750

source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

IGIN

Query Match 23.4%; Score 491.6; DB 12; Length 750;
Best Local Similarity 89.3%; Pred. No. 1e-98;
Matches 598; Conservative 0; Mismatches 64; Indels 8; Gaps 6;
1098 CACAGACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCCTGG- 1156
|||
80 CAGAACTCATAGATTACTGCAAGATTCCCTCTGGACCTGGAGAAGTTTCTTCTCTGGA 139
|||
1157 CTTACAGAAGCTGAAAC--AACTGCCAATGTCTCTACAGGATGCTACCCG-TAAGGAAAG 1213

Db 140 TTTACGGAAGCAGAAAAACAGAGCTGCCAATGTCTTACAGGACGCTTCCGCTAAGGAGAAG 199
|||||
1214 CTCCTAGAAGACTCCAAGGG--AGTAAAAAGAGCTGATGAACAATGGCAAGACCTCCAAG 1271
|||||
200 CTCCTAGCAAGACTCCAGGGGCAGTCAGAGAGCTGATGAACAATGGCAAGATCTCCAAG 259
|||||
1272 GTGAAA-TTGAAAGCTCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAAATC 1330
|||||
260 GAGAAAGTTGAAACTCACACAGATATCTATCACAATCTTGATGAAAAATGGCCAAAAAATC 319
|||||
1331 CTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATG 1390
|||||
320 CTGAGATCCCTGGAAGGTTCCGATGATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATG 379
|||||
1391 AACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGGAAGCC 1450
|||||
380 AATTTCAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCA 439
|||||
1451 AGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAG 1510
|||||
440 AGTTCTGACCAAGTGAAGCGTTTGCATCTTCTCTTCAGGAACCTTCTTGTGCTGCTACAG 499
|||||
1511 CTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCCAGCAGTTTCAG 1570
|||||
500 CTGAAAGATGATGAACCTGAGCCGTCAGGCACCCATCGGTGCTGATTTCCAGCAGTTTCAG 559
|||||
1571 AAGCA-GAAGCATGTACATAGGSCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAAT 1629
|||||
560 AAGCAGCAATGATATACATAGGSCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAAT 619
|||||
1630 CATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGA 1689
|||||
620 CATGAGTACTCTTGAGACTGTGAGATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGA 679
|||||
1690 GAAACTCTACAGAGCCCGCAGAGAGCTGCCTCTCTGAGGAGAGAGCCCGAGAAATGTCACTCG 1749
|||||
680 GAAACTCTACAGAGCCCGCAGAGAGACTGCCTCTCTGAGGAGAGAGCTCAGAAATGTCACTCG 739
|||||
1750 GCTTCTACGA 1759
|||||
740 GCTCCTACGA 749
|||||

RESULT 11
AL121550
LOCUS
DEFINITION
DKFZp762L078 r1 762 (synonym: hm12) Homo sapiens cDNA clone
579 bp mRNA linear EST 04-SEP-2003
DKFZp762L078 5', mRNA sequence.
AL121550
AL121550.1 GI:5927551
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 579)
Ottewaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Ottewaelder, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No si sequence
available.
This clone (DKFZp762L078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

b 540 TGAGTGGGATAAGCTAAATCTACGTTCTGTGATTGGCAAAAGAGATAGATGATGCTCT 599
Y 1843 TGAAGACTCCAGGAACCTTCAAGAGGCCCGGATGAGCTGGACCTCAAGCTGCGCCAAGC 1902
b 600 TGAAGAGACTGCAGGGTCTTTCAGGAGGCAATGATGAACCTNNNCTGAAACTGCGCCAGGC 659
Y 1903 TGA 1905
b 660 TGA 662

ESULT 13
U313510
OCUS
EFINITION BU313510 644 bp mRNA linear EST 28-NOV-2002
603540290F1 CSEQCHN61 Gallus gallus cdna clone ChEST508c24 5', mRNA
sequence.
CCENSION BU313510
ERSION BU313510.1 GI:25821511
EYWORDS EST.
OURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 644)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

22335534
12445392
PUBMED
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

1. .644
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="ChEST508c24"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN61"
/note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

RIGIN
Query Match 18.5%; Score 389; DB 13; Length 644;
Best Local Similarity 75.5%; Pred. No. 7e-76;
Matches 482; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
Y 1449 CCAGTTCTGACCAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGGCTAC 1508

Db 7 CGAGCACAGACCAGTGGAGCGTTTACATCTCTCTTTCAGGAACCTTTGGCATGGCTGC 66
QY 1509 AGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTC 1568
Db 67 AATTGAAGGAGGATGAATTAATAACAGCAAGCACCATTGGTGAGATATCCCACTGTGC 126
QY 1569 AGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTAAAGAACCTGTAA 1628
Db 127 AGAAGCAGAAATGATGTTTCATAGGACTTTCAAGAGGGAGCTGAAAACAAAAGAACCTGTTA 186
QY 1629 TCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAG 1688
Db 187 TCATGAATGCACTTGAGACTGTGCGACTCTTCTGGCAGATCAACCAGTAGAGGACTGG 246
QY 1689 AGAAACTCTACCAGGAGCCGACAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAATGTCACTC 1748
Db 247 AAAAGTCTATCCAGAACCAAGAGACCTATCACCCTGAGGAGAGGGGCCCAAGATGTCACTA 306
QY 1749 GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACT 1808
Db 307 AAGTTCTCCGAGGCAAGCAGATGATGTCAGAACTGAGTGGGATAAGCTAAATCTACGTT 366
QY 1809 CCGTGACTGGCAGAGAAAAATAGATGAGAGCCCTTGAAAGACTCCAGGAACCTTCAAGAGG 1868
Db 367 CTGCTGATTGGCAAAAGAGATAGATGATGCTCTTGAAAGACTGCAGGGTCTTCAGGAGG 426
QY 1869 CCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGC 1928
Db 427 CAATGATGAACCTAGACCTGAAACTGCGCCAGGCTGAAGCATTCAGGGATCCTGGCAGC 486
QY 1929 CCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTC 1988
Db 487 CAGTGGGGGATCTGCTGATAGACTCTCTGCGAGATCACTTAGAAAAAGTCAAGGTTTATC 546
QY 1989 GAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGCCAGC 2048
Db 547 GAGCAGAAATGGTGGCCCTTANAGAGAAAGGTGATCAAGTCAATGAGCTGGCTCACCGGT 606
QY 2049 TTACCACCTTTGGGCATTCAGCTCTCACCGTATACCTC 2086
Db 607 TCGTCCCTCCCTGATATTTCAGTTCTCCCCATACACTCTC 644

RESULT 14
BC011062
LOCUS BC011062 2334 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, Similar to utrophin, clone IMAGE:3979320, mRNA.
ACCESSION BC011062
VERSION BC011062.1 GI:15029695
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 2334)
Strausberg,R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

REMARK
COMMENT
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES

source
1. .2334
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979320"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 16.3%; Score 343; DB 11; Length 2334;
Best Local Similarity 59.0%; Pred. No. 1.6e-65;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

2y 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTA 163
db 1042 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 1101

2y 164 TTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAATGAT 223
db 1102 CTGACGTGGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1161

2y 224 GTGGAAGTGTGAAAGACCAGTTTTCATCTACTCATGAGGGTACATGATGGATTGACAGCC 283
db 1162 GTCGAAGAAGTCAAAAGAGCAGTTTGTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1221

2y 284 CATCAGGGCCGGTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGACAGGAAAA 343
db 1222 CACCAGAGCAGCTGGGGAGCGCTCTCGAGGCTGGCAACGAGCTGATGACACAAAGGACT 1281

Qy 344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
Db 1282 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGCTGAATGCAAGGTGG 1341

Qy 404 GAATGCCCTCAGGTAGCTAGCATGTGAAAAACAAAGCAATTACATAGAGTTTAAATGGAT 463
Db 1342 GAGGCGCTCCGGTGGAGAGCATGGAGAGGCAGTCCCGGTGCACGCTCTGTATGGAG 1401

Qy 464 CTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACA 523
Db 1402 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCAGAAAGCTGTTCAA 1461

Qy 524 AGGAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAA 583
Db 1462 CAGAAGATGGAGAGCCTCCCGCTGGGTGATGACCTGCCCTCCTGCAGAAAGCTGTTCAA 1521

Qy 584 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGACAAGTCAAGGTCAATTCTCTCACT 643
Db 1522 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAAC 1581

Qy 644 CACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGGAACCTGCTGCTTTGGAAGAA 703
Db 1582 CACATGGTGGTGTAGTTGATGAATGAAACAGTGGGAGAGTGCACAGCTCTTCTGGAAGAT 1641

Qy 704 CAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGG 763
Db 1642 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 1701

Qy 764 GTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTT 823

Db 1702 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTCTGTG 1761

Qy 824 AGTGCATGGCTTTCAGAAAAAGAAAGATGCGAGTGAAACAAGATTCACACAACACTGGCTTTAAA 883
Db 1762 GAGGCTTGGCTCACCGAAAAGGAAGAGCGCTTTGAATAAAGTTCAAAACCCAGCAACTTTAAA 1821

Qy 884 GATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAA 943
Db 1822 GACCAGAAGGAACATAAGTGTGAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1881

Qy 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCTCAAAACAAGATCTTCTTTCAACACTG 1003
Db 1882 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGCCCAATTACTC 1941

Qy 1004 AAGAATAAGTCAGTGACCCAGAAAGCCGAAGCATGGCTGGATAACTTTGCCCGGTGTGG 1063
Db 1942 AGTAATCCCAAGGCATCTAAGAAGATCAACAGTGAAGTCTCTGAGGAGCTAACACAGAGATGG 2001

Qy 1064 GATAATTTAGTCCAAAAAATCTTGAAAAGAGTACAGCACAG 1102
Db 2002 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTTAACCAG 2040

RESULT 15
AY407022
LOCUS AY407022 9915 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens UTRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY407022
VERSION AY407022.1 GI:39762993
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9915)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5552), 1960-1963 (2003)
PUBMED 14571302
REFERENCE 2 (bases 1 to 9915)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
1..9915
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>9915
/gene="UTRN"
/locus_tag="HCM2752"

ORIGIN
Query Match 15.9%; Score 334.6; DB 29; Length 9915;
Best Local Similarity 62.5%; Pred. No. 1.5e-63;
Matches 523; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

Qy 103 CAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAAGT 162
Db 918 CAGCACTGTCACTGAGGTTGACATGGATCTGGACAGCTATCAGATTGCGTTGGAGGAAGT 977

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

4 nucleic - nucleic search, using sw model

run on: April 4, 2004, 16:58:13 ; Search time 101.891 Seconds
(without alignments)
11443.127 Million cell updates/sec

itle: US-09-845-416-2_COPY_900_3000

erfect score: 2101
equence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

earched: 682709 seqs, 277475446 residues

otal number of hits satisfying chosen parameters: 1365418

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	1131	53.8	5952	4	US-09-687-875A-1
2	1103.4	52.5	13977	4	US-09-484-970B-60
c 3	866.6	41.2	19307	3	US-08-836-022A-10
c 4	866.6	41.2	19307	3	US-09-427-048A-10
5	343	16.3	6045	4	US-09-091-501B-7
6	343	16.3	10320	4	US-09-091-501B-9
7	200.8	9.6	3915	4	US-09-976-594-93
8	79.4	3.8	200	4	US-09-091-501B-5
9	78.6	3.7	200	4	US-09-091-501B-4
10	78.6	3.7	200	4	US-09-091-501B-6
c 11	75.6	3.6	7218	1	US-08-232-463-14
12	63.6	3.0	238	4	US-09-687-875A-13
c 13	46.2	2.2	505	4	US-09-621-976-15639
c 14	44	2.1	1230025	4	US-09-198-452A-1
15	43.4	2.1	1179	4	US-09-107-532A-1186
16	40.4	1.9	832	4	US-09-621-976-2813
17	40.4	1.9	2223	1	US-08-257-073-4
c 18	39.4	1.9	193303	4	US-09-497-855A-37
c 19	39.4	1.9	193303	4	US-09-497-855A-44
20	39.2	1.9	16995	4	US-08-961-527-82
21	38.6	1.8	1751	4	US-09-620-312D-847
22	38.6	1.8	1995	1	US-08-425-069-3
23	38.6	1.8	1995	2	US-08-317-844B-3
24	38.2	1.8	1131	6	5180810-3
25	38.2	1.8	1784	6	5180810-2
26	38.2	1.8	1848	4	US-09-134-001C-447
27	38.2	1.8	4929	4	US-09-620-312D-674

c 28	38	1.8	1394	4	US-09-247-155-76	Sequence 76, Appl
c 29	37.4	1.8	2082	3	US-08-985-335-4	Sequence 4, Appli
c 30	37.4	1.8	2082	3	US-09-410-372-4	Sequence 4, Appli
c 31	37.2	1.8	2160	4	US-09-092-218-1	Sequence 1, Appli
c 32	37.2	1.8	2915	4	US-09-336-115C-5	Sequence 5, Appli
c 33	37.2	1.8	3902	4	US-08-961-527-212	Sequence 212, App
34	36.8	1.8	1886	6	5210183-1	Patent No. 5210183
35	36.6	1.7	1845	4	US-08-887-534A-22	Sequence 22, Appl
36	36.6	1.7	1845	4	US-09-527-431-22	Sequence 22, Appl
c 37	36.6	1.7	7075	4	US-08-956-171E-263	Sequence 263, App
38	36.2	1.7	289	3	US-09-007-005-17	Sequence 17, Appl
39	36.2	1.7	289	3	US-09-244-796-17	Sequence 17, Appl
40	36.2	1.7	2447	2	US-09-014-969-14	Sequence 14, Appl
c 41	36.2	1.7	11049	4	US-10-204-708-23	Sequence 23, Appl
42	36.2	1.7	168575	4	US-09-426-290-1	Sequence 1, Appli
43	36	1.7	608	3	US-09-385-982-236	Sequence 236, App
c 44	36	1.7	724	4	US-08-956-171E-832	Sequence 832, App
45	36	1.7	2873	4	US-08-630-915A-193	Sequence 193, App

ALIGNMENTS

RESULT 1

US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match	53.8%;	Score 1131;	DB 4;	Length 5952;
Best Local Similarity	68.6%;	Pred. No. 0;		
Matches 2101;	Conservative	0;	Mismatches	0; Indels 960; Gaps 1;
QY	1	GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGCCCAT	60	
Db	891	GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGCCCAT	950	
QY	61	TCCTTCACAGCATTTCGAAGCTCCTGAAGACAAGTCATTTCGACGATTCATTGATGGAGAG	120	
Db	951	TCCTTCACAGCATTTCGAAGCTCCTGAAGACAAGTCATTTCGACGATTCATTGATGGAGAG	1010	
QY	121	TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC	180	
Db	1011	TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC	1070	
QY	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240	
Db	1071	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1130	
QY	241	CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTTGG	300	

1131 CCAGTTTCATACTCATGAGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG 1190
301 TAATATTCATAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
1191 TAATATTCATAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1250
361 AACTGAAGTACAAGAGCGAGATGGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
1251 AACTGAAGTACAAGAGCGAGATGGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 1310
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
1311 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1370
481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 540
1371 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 1430
541 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAAGGTGCTTCA 600
1431 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAAGGTGCTTCA 1490
601 AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTCCTCACTCACATGGTGGTGTAGT 660
1491 AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTCCTCACTCACATGGTGGTGTAGT 1550
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 720
1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 1610
721 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAGACAT 780
1611 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAGACAT 1670
781 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTTAGTGATGGCTTTTCAGA 840
1671 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTTAGTGATGGCTTTTCAGA 1730
841 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
1731 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1790
901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
1791 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1850
961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGTAAGATAAGTCAGTGAC 1020
1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGTAAGATAAGTCAGTGAC 1910
1021 CCAGAGACGGAAGCATGGCTGATTAACCTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
1911 CCAGAGACGGAAGCATGGCTGATTAACCTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1970
1081 ACTTGAAAGAGTACAGCACAG----- 1102
1971 ACTTGAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAAACCGGATGT 2030
1103 ----- 1102
2031 GGAAGAGATTTTGTCTAAAGGGCAGCATTTGTACAGGAAAAACAGCCACTCAGCCAGT 2090
1103 ----- 1102
2091 GAAGAGGAGTTAGAAGATCTGAGCTCTGAGTGGAGCGGTAAACCGGTTTACTTCAAGA 2150
1103 ----- 1102
2151 GCTGAGGGCAAGCAGCCTGACCTAGCTCCTGGACTGACCACACTATTGGAGCCTCTCCTAC 2210
1103 ----- 1102
2211 TCAGACTGTACTCTGGTGACACAACCTGTGGTTACTTAAGGAAACTGCCATCTCCAAACT 2270

QY 1103 ----- 1102
Db 2271 AGAAATGCCATCTTCTTGATGTGGAGGTACCTGCTCTGGCAGATTTCAACCGGCTTG 2330
QY 1103 ----- 1102
Db 2331 GACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTATAAAATCACAGAGGGTGAT 2390
QY 1103 ----- 1102
Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGA 2450
QY 1103 ----- 1102
Db 2451 TTTGGAACAGAGGCGTCCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAA 2510
QY 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAACAAATCATTTACGGATCGAATTGAAAGAAATTCAGAA 2570
QY 1103 ----- 1102
Db 2571 TCAGTGGGATGAAGTACAAGAAACACCTTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630
QY 1103 ----- 1102
Db 2631 AAAGGATTCAACACAAATGGCTGGAAGCTAAGGAAGAACTGAGCAGGTCTTTAGGACAGGC 2690
QY 1103 ----- 1102
Db 2691 CAGAGCCAAAGCTTGAGTCATGGAAGGAGGGTCCCTATACAGTAGATGCAATCCAAAAAGAA 2750
QY 1103 ----- 1102
Db 2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAATGTAGATGT 2810
QY 1103 ----- 1102
Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACCAGAAAAAGT 2870
QY 1103 ----- 1102
Db 2871 CCACATGATAACAGAGAAATATCAATGCCTCTTGGAGAAGCATTCATAAAAGGGTGAGTGA 2930
QY 1103 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
Db 2931 GCGAGAGGTGCTTTTGGAAAGAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2990
QY 1141 AAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAAATGTCTCTACGATGCTAC 1200
Db 2991 AAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAAATGTCTCTACGATGCTAC 3050
QY 1201 CCGTAAGGAAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 1260
Db 3051 CCGTAAGGAAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 3110
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320
Db 3111 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 3170
QY 1321 CCAAAAAATCCTGAGATCCCTGGNAGGTTCCGATGATGAGTCTCTGTTTACAAGACGTTT 1380
Db 3171 CCAAAAAATCCTGAGATCCCTGGNAGGTTCCGATGATGAGTCTCTGTTTACAAGACGTTT 3230
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAANAAGTCTCTCAACATTTAGGTCCCA 1440
Db 3231 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAANAAGTCTCTCAACATTTAGGTCCCA 3290
QY 1441 TTTTGAAGCCAGTTCTGACCAGTGAAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 3291 TTTTGAAGCCAGTTCTGACCAGTGAAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 3350

1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3410
1561 AGCAGTTCAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGA 1620
3411 AGCAGTTCAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGA 3470
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGA 1680
3471 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGA 3530
1681 AGGACTAGAGAAACTCTACCAGGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGAA 1740
3531 AGGACTAGAGAAACTCTACCAGGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGAA 3590
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800
3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 3650
1801 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGAGACCCCTTGAAAGACTCCAGGA 1860
3651 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGAGACCCCTTGAAAGACTCCAGGA 3710
1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCAAGCTGAGGTGATCAAGGGATC 1920
3711 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCAAGCTGAGGTGATCAAGGGATC 3770
1921 CTGGCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
3771 CTGGCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3830
1981 GGCACCTTCAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 2040
3831 GGCACCTTCAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 3890
2041 TCGCCAGCTTACCACCTTTGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
3891 TCGCCAGCTTACCACCTTTGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3950
2101 C 2101
3951 C 3951

3-09-484-970B-60
Sequence 60, Application US/09484970B
Patent No. 6426186

GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmuth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 60
LENGTH: 13977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
NAME/KEY: unsure
LOCATION: 11721-11761, 12294, 13969
OTHER INFORMATION: a, t, c, g, or other

3-09-484-970B-60

Query Match 52.5%; Score 1103.4; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAT 1158
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGSCAGTTCAATGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGSCAGTTCAATGATGGAGAG 1218
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1278
Qy 181 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
Qy 241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1398
Qy 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
Qy 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
Qy 481 GAAAGAGTTGAATGAATGGCTAACAAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGAATGGCTAACAAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 1638
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA 1698
Qy 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1758
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1818
Qy 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
Qy 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 1879 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
Qy 841 AAAAGAAGATGCAGTGAAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAATGTT 1998
Qy 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
Qy 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAA 1080
Db 2119 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAA 2178

2Y 1081 ACTTGAAGAAGAGTACAGCACAGACT 1105
13328 ACAGTTTTCATGCTCATGAGGGATTTCATGATGGATCTGTCATCTCATCAAGGACATTGTTGG 13269
Db 2179 ACTTGAAGAAGAGTACAGCACAGATT 2203

RESULT 3

JS-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A

FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVEN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

JS-08-836-022A-10

Query Match 41.2%; Score 866.6; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 4.1e-254;
Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

2Y 1 GAGCTATGCTTACACACAGGCTGCTTATGTACCACTCTGTGACCTTACAGGAGCCCAT 60
Db 13568 GAGTTATGCTTACACAGGCTGCTTATGTGCCCACCTCTGATTCACACAGAGCCCTA 13509

2Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGTATGGAGAG 120
Db 13508 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGTATGGAGAG 13449

2Y 121 TGAAGTAACTGGACCGCTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 13448 GGAAGTAACTGGATAGTTACCAAACTGCTTTAGAGAAGTACTTTTCATGGCTTCTTTC 13389

2Y 181 TGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTTAATGATGTGGAATGGTGAAGA 240
Db 13388 TGCCGAGGATACATTCGAGCACAGGAGAGATTTCAAAATGATGTTGAAGAAGTGAAGA 13329

QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG. 300
Db 13328 ACAGTTTTCATGCTCATGAGGGATTTCATGATGGATCTGTCATCTCATCAAGGACATTGTTGG 13269
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 13268 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAAGGAAAAATTATCAGAAGATGAAGA 13209
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 13208 AGCTGAAGTGCAAGAACAATGAATCTCCTAAATTCAGATGGGAATGTCTCAGGGTAGC 13149
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 480
Db 13148 TAGCATGGAACAAACAAAGCAATTTACCAAAAGTTTCTAATGATCTCCAGAATCAGAAATT 13089
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 13088 AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 13029
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTCCTTCA 600
Db 13028 GCCCTTTGGACCTGATCTTGAAGATCTAAATGCCAAGTACAAACAATAGGTCCTTCA 12969
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
Db 12968 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 12909
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 12908 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 12849
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 780
Db 12848 AGATCGATGGCAAAATATCTGCAGATGGACTGAAGACCCGCTGGATTGTTTACAAGATAT 12789
QY 781 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 12788 TCTTCTAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 12729
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 900
Db 12728 AAAAGAAGATGCATGAAGAACAATTCAGACAAGTGGCTTTAAAGATCAAAAATGAAATGAT 12669
QY 901 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 12668 GTCAAGTCTTCAAAAAATATCTACTTTAAAAATAGATCTAGAAAAAGAAAAAGCAATCCAT 12609
QY 961 GGGCAAACTGATTCATCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 12608 GGAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAAAATAAGTCAGTGAC 12549
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 12548 TCAAAAGATGGAATACTGGATGGAAAACTTTGCACACAGCTTGGGACAATTTAACCCAAAA 12489
QY 1081 ACTTGAAAAAGAGTACAGCAGACT 1105
Db 12488 ACTTGAAAAAGAGTTCAGCACAAAT 12464

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weitzman, Matthew

TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVN.008PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

S-09-427-048A-10

Query Match 41.2%; Score 866.6; DB 3; Length 19307;

Best Local Similarity 86.5%; Pred. No. 4.1e-254;

Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

y	1	GAGCTATGCCTACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT	60
b	13568	GAGTTATGCCCTTCACAGGCTGCTTATGTTGCCACCTCTGATTCACACAGAGCCCTA	13509
y	61	TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG	120
b	13508	TCCTTCACAGCATTTGGAAGCTCCAGAGACAAAGTCATTGACAGTTTCATTGATGGAGAC	13449
y	121	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAAGTATATCGTGGCTTCCTTC	180
b	13448	GGAAGTAAATCTGGATAGTTACCAACTGCTTTAGAAGAAAGTATTCATGGCTTCCTTC	13389
y	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA	240
b	13388	TGCCGAGGATACATTGCGAGCACAAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA	13329
y	241	CCAGTTTCATCTCATGGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG	300
b	13328	ACAGTTTCATGCTCATGGGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGG	13269
y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATATCAGAAGATGAAGA	360
b	13268	TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAATATCAGAAGATGAAGA	13209
y	361	AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC	420
b	13208	AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAGATGGGAATGCTCTCAGGGTAGC	13149
y	421	TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAACT	480
b	13148	TAGCATGAAAAACAAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAATCAGAAAT	13089

QY	481	GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA	540
Db	13088	AAAAGAACTAGATGACTGGTTAAACAAAACCTGAAGAGAGAACTAAAGAAAATGGAGGAAGA	13029
QY	541	GCCTCTTGGACCTGATCTTGAAGACCTAAACGCAAGTACAAACAACTAAAGGTGCTTCA	600
Db	13028	GCCCTTTGGACCTGATCTTGAAGATCTAAATGCAAGTACAAACAACTAAAGGTGCTTCA	12969
QY	601	AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATCTCTCACTCACATGGTGGTGTAGT	660
Db	12968	AGAAGATCTAGAACAGGAGCAGGTCAAGGTCACTCGCTCACTCACATGGTGTAGTGGT	12909
QY	661	TGATGAATCTAGTGAGATCACGCAACTGCTGCTTGGGAAGCAACAACTTAAGGTATTGGG	720
Db	12908	TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTGGGAAGCAACAACTTAAGGTACTGGG	12849
QY	721	AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT	780
Db	12848	AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT	12789
QY	781	CCTTCTCAAATGGCAACGTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA	840
Db	12788	TCCTTCAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA	12729
QY	841	AAAAGAAGATGCAGTGAACAAAGATTTCACACAACTGGCTTTAAAGATCAAAAATGAAATGT	900
Db	12728	AAAAGAAGATGCAATGAAGAACATTTCAGACAACTGGCTTTAAAGATCAAAAATGAAATGT	12669
QY	901	ATCAAGTCTTCAAAAACCTGCGCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
Db	12668	GTCAAGTCTTCAAAAATATCTACTTTTAAATATAGATCTAGAAAAGAAAAGCAATCCAT	12609
QY	961	GGGCAAACTGTATTCACCTCAAAACAAAGATCTTCTTCAACACTGAAGHATAAGTCAGTGAC	1020
Db	12608	GGAAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCAGCTGAAAAATAAGTCAGTGAC	12549
QY	1021	CCAGAAGACGGGAAGCATGGCTGGATAACTTTTCCCGGTGTTGGGATAATTTAGTCCAAA	1080
Db	12548	TCAAAGATGGAAATCTGGATGGAAAACCTTTGCACAACGTTGGGACAATTTAACCCAAA	12489
QY	1081	ACTTGAAAAAGAGTACAGCACAGACT 1105	
Db	12488	ACTTGAAAAAGAGTTTCAGCACAAAT 12464	

RESULT 5

US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Urothrin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
JS-09-091-501B-7

Query Match 16.3%; Score 343; DB 4; Length 6045;
Best Local Similarity 59.0%; Pred. No. 4.5e-94;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

2Y 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
Db 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
2Y 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973
Db 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973
2Y 164 TTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGAT 223
Db 164 TTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGAT 223
2Y 974 CTGACGTGGCTGTCTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033
Db 974 CTGACGTGGCTGTCTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033
2Y 224 GTGGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACAGCC 283
Db 224 GTGGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACAGCC 283
2Y 1034 GTCGAAGAAGTCAAGAGAGCAGTTTGTCTACCCATGAACTTTTATGATGAGCTGACAGCA 1093
Db 1034 GTCGAAGAAGTCAAGAGAGCAGTTTGTCTACCCATGAACTTTTATGATGAGCTGACAGCA 1093
2Y 284 CATCAGGSCGGTGGTAAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAA 343
Db 284 CATCAGGSCGGTGGTAAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAA 343
2Y 1094 CACCAGAGCAGCTGGGAGCGTCTGCAGGCTGGCAACCAGCTGATGACACAAGGGACT 1153
Db 1094 CACCAGAGCAGCTGGGAGCGTCTGCAGGCTGGCAACCAGCTGATGACACAAGGGACT 1153
2Y 344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGG 403
Db 344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGG 403
2Y 1154 CTGTCAGAGGAGGAGGATTTGAGATCCAGGAACAGATGACCTTGCTGATGCAAGGTGG 1213
Db 1154 CTGTCAGAGGAGGAGGATTTGAGATCCAGGAACAGATGACCTTGCTGATGCAAGGTGG 1213
2Y 404 GAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
Db 404 GAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
2Y 1214 GAGGCGTCCGGTGGAGAGCATGGAGAGGCAGTCCCGGTGCACGACGCTCTGATGGAG 1273
Db 1214 GAGGCGTCCGGTGGAGAGCATGGAGAGGCAGTCCCGGTGCACGACGCTCTGATGGAG 1273
2Y 464 CTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAGACA 523
Db 464 CTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAGACA 523
2Y 1274 CTGCAGAAACAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAAGCGCAT 1333
Db 1274 CTGCAGAAACAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAAGCGCAT 1333
2Y 524 AGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTAAACCCCAAGTACAA 583
Db 524 AGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTAAACCCCAAGTACAA 583
2Y 1334 CAGAAGATGGAGAGCCCTCCGCTGGGTGATGACCTTGCCCTCCCTGCAGAACTGCTTCAA 1393
Db 1334 CAGAAGATGGAGAGCCCTCCGCTGGGTGATGACCTTGCCCTCCCTGCAGAACTGCTTCAA 1393
2Y 584 CAACATAACGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGTCAATTTCTCTCACT 643
Db 584 CAACATAACGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGTCAATTTCTCTCACT 643
2Y 1394 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAAT 1453
Db 1394 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAAT 1453
2Y 644 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAAGAA 703
Db 644 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAAGAA 703
2Y 1454 CACATGGTGGTGGTAGTTGATGAATGAAGACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1513
Db 1454 CACATGGTGGTGGTAGTTGATGAATGAAGACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1513
2Y 704 CAACCTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGG 763
Db 704 CAACCTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGG 763
2Y 1514 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGGAACACGTTGG 1573
Db 1514 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGGAACACGTTGG 1573
2Y 764 GTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTT 823
Db 764 GTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTT 823
2Y 1574 AACAGGTTGCAAGAAATCAGTATTTCTGTGGCAGGAATTAATTGGAAGAGCAGTGTCTGTTG 1633
Db 1574 AACAGGTTGCAAGAAATCAGTATTTCTGTGGCAGGAATTAATTGGAAGAGCAGTGTCTGTTG 1633
2Y 824 AGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTAAA 883
Db 824 AGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTAAA 883
2Y 1634 GAGGCTTGGCTCACCGAAAGGAAGAGGCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693
Db 1634 GAGGCTTGGCTCACCGAAAGGAAGAGGCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693
2Y 884 GATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCGGTTTTAAAAGCGGATCTAGAA 943
Db 884 GATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCGGTTTTAAAAGCGGATCTAGAA 943
2Y 1694 GACCAGAAGGAACCTAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1753
Db 1694 GACCAGAAGGAACCTAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1753
2Y 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTG 1003
Db 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTG 1003
2Y 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGCCAAATTAATC 1813
Db 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGCCAAATTAATC 1813
2Y 1004 AAGAATAAGTCACTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTTGG 1063
Db 1004 AAGAATAAGTCACTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTTGG 1063

Db 1814 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGAATCTGTAGGAGCTAACAACAGAGATGG 1873
QY 1064 GATAATTTAGTCCAAAAAAGTGTGAAAAAGAGTACAGCACAG 1102
Db 1874 GATTCTCTGTTTCAGAGACTCGAAGACTCTTTCTAACCAG 1912

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Urophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 16.3%; Score 343; DB 4; Length 10320;
Best Local Similarity 59.0%; Pred. No. 6.6e-94;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

QY 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
Db 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973
QY 164 TTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGAT 223
Db 974 CTGACGTGGCTGTCTCCGCGGAGGACAGTTCACCCATGAAACTTTATGATGGAGCTGACAGCA 1093
QY 224 GTGGAAGTGGTGAAGAACCCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCC 283
Db 1034 GTCGAAGAAGTCAAAAGAGCAGTTTGTCTACCCATGAAACTTTATGATGGAGCTGACAGCA 1093
QY 284 CATCAGGCGCGGTTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAA 343
Db 1094 CACCAGAGCAGCTGGGAGCGTCTCTGAGGCTGGCAACCAGCTGATGACACAAGGGAAT 1153
QY 344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGG 403
Db 1154 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGTGAATGCAAGGTGG 1213
QY 404 GAATGCCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGAT 463
Db 1214 GAGGCGCTCCGGGTGGAGAGCATGGAGAGGACGTCCTCCGGCTGCACGACGCTCTGATGGAG 1273

Y 464 CTCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACA 523
b 1274 CTGCAGAAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAAGAGCGCAT 1333
Y 524 AGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAA 583
b 1334 CAGAAGATGGAGAGCCCTCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAAGCTGCTTCAA 1393
Y 584 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTACGGTCAATTTCTCTCACT 643
b 1394 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTTCCTTAACT 1453
Y 644 CACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTTGGAGAA 703
b 1454 CACATGGTGTGATTTGTGATGAAAAACAGTGGGAGAGTGCCACAGCTCTTCTTGGAAAT 1513
Y 704 CAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGAACAGAACCCGTGG 763
b 1514 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGTGGAAGTGAAGAACGTTGG 1573
Y 764 GTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTTCTTACTGAAGAAACAGTGCCTTTT 823
b 1574 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATATTGGAAGAGCAGTGTCTGTTG 1633
Y 824 AGTGCATGGCTTTCAGAAAAAGAAAGATGCAAGTGAACAAAGATTTCACAAACTGGCTTTAAA 883
b 1634 GAGGCTTGGCTCACCCGAAAGGAAGAGGCTTTGAATAAAGTTCAAAACAGCAACTTTAAA 1693
Y 884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAA 943
b 1694 GACCAGAAAGAACTAAGTGTCAAGTGTCCGGGCTGGCTATATTGAAGGAAGACATGGAA 1753
Y 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAAGATCTTCTTCAACACTG 1003
b 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTAATC 1813
Y 1004 AAGAAATAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGG 1063
b 1814 ACTAATCCCAAGGCATCTAAGAAGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 1873
Y 1064 GATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAG 1102
b 1874 GATTCTCTGTTTCAGAGACTCGAAGACTCTTTCTAACCCAG 1912

ESULT 7

S-09-976-594-93
Sequence 93, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 93
LENGTH: 3915
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 290344.1
S-09-976-594-93

Query Match 9.6%; Score 200.8; DB 4; Length 3915;
Best Local Similarity 56.0%; Pred. No. 1.2e-50;
Matches 400; Conservative 0; Mismatches 312; Indels 2; Gaps 1;

QY 1387 CATGAACCTCAAGTGGAGTGAACCTTCGGAATAAAGTCTCTCAACATTAGTCCCATTTGGA 1446
Db 544 CATGAATCTGTGTTGGATGAATAAATAAATAAAGTCTCAACCTCCGCGCTCGCCTAGA 603
QY 1447 AGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCT 1506
Db 604 GGCTTCTCTCAGACCAACAGTGGAAAGCTTCAAGTCCCTCTTCAAGAGATATTGACTGGCT 663
QY 1507 ACAGCTGAAAGATGATGAATTAAGCCGGCAGGACACCTATTGGAGGCGGACTTTCCAGCAGT 1566
Db 664 CAGCCAAAAGGATGAGGAGTTGTAGCTCAGCTCAGCTGCCCCCTACAGGGGGATGTGGCCCTGGT 723
QY 1567 TCAGAACGAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGT 1626
Db 724 GCAACAGGAGAAAGGAGACACATGCGGCCCTTTATGGAAGAAGTCAAGTCTCGGGGCCCTTA 783
QY 1627 AATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACT 1686
Db 784 CATCTATTCTGTCTGGAGTCAAGTCAAGGCTTCTGTCTCCAGCACCCATTTGAGGAGTT 843
QY 1687 AGAGAAACTCTACAGGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCACAGAAATGTCAAC 1746
Db 844 AGAGAGCCTCATTTCTGAGAGCAAAAGATACCTCCCGAAACAGCGGATCCAGAAATCTCAG 903
QY 1747 TCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAAACCTTGCA 1806
Db 904 CCGCTTTGTATGGAAGCAGCGCACGTTGGCCAGTGAACCTGTGGGAGAGTTGACAGCCCG 963
QY 1807 CTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGA 1866
Db 964 CTGTGTGGACCAAGCACCCGTACATTTGAGCGGACTCTGGAGCAGCTCTTGGAGATTCAAGG 1023
QY 1867 GGCCAGGATGAGCTGGACCTCAAGCTCGCCAGCTGAGGTGATCAAGGGATCCTGGCA 1926
Db 1024 G--CATGGAGGAACTAAGCACTACTCTGAGCCAAGCTGAGGGAGTCCGAGCCACTTGGGA 1081
QY 1927 GCGCGTGGCGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACT 1986
Db 1082 GCGCATTTGGGATCTCTTCTATTGATTCACTCCAGAGACACATCCAGGCTATTAAAGTGT 1141
QY 1987 TCGAGGAGAAATTTGGCCCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGTCTCGCCA 2046
Db 1142 CAAAGAAGAATTTCTCCCATGAAAGATGAGTAAAGTTGGTGAATGATCTGGCCACCA 1201
QY 2047 GCTTACCACCTTTGGGCATTCAGCTCTCACCCGATAAACCTCAGCAGCTCTGGAAGA 2100
Db 1202 ACTTCCCATTTCTGATGTGCACTTGTCAATGAGAGATTTCCAGGGCCCTTGAACA 1255

RESULT 8

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9pt-F1s
S-08-232-463-14

Query Match 3.6%; Score 76.6; DB 1; Length 7218;
Best Local Similarity 6.7%; Pred. No. 2.1e-12;
Matches 28; Conservative 236; Mismatches 155; Indels 0; Gaps 0;
Y 213 TTCTAATGATGGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGG 272
b 1474 TATCTATGCAAGTAGTAAAGAGATAGAGAATTGGTACRRRRRRRRRRRRRRRR 1415
Y 273 ATTTGACAGCCCATCAGGGCCGGTGTGTAATATTCTACAATTGGGAAGTAAGCTGATG 332
b 1414 RRR 1355
Y 333 GAACAGGAAATATATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAA 392
b 1354 RRR 1295
Y 393 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAG 452
b 1294 RRR 1235
Y 453 TTTTAATGGATCTCCAGAATCAGAAATCAGAAAGTTGAATGACTGGCTAAACAAAACAG 512
b 1234 RRR 1175
Y 513 AAGAAAGAACAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAC 572
b 1174 RRR 1115
Y 573 GCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTC 631
b 1114 RRR 1056

ESULT 12

S-09-687-875A-13
Sequence 13, Application US/09687875A
Patent No. 6544786
GENERAL INFORMATION:
APPLICANT: Xiao, Xiao
APPLICANT: Liu, Paul
TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPliced PE
FILE REFERENCE: 00792
CURRENT APPLICATION NUMBER: US/09/687,875A
CURRENT FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/158,868
PRIOR FILING DATE: 1999-10-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 238

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match 3.0%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.9e-09;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1348 TTCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGA 1407
Db 169 TTCGACGACGACGAGTACTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGA 228
QY 1408 ACTTCGAAA 1417
Db 229 ACTTCGAAA 238

RESULT 13

US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match 2.2%; Score 46.2; DB 4; Length 505;
Best Local Similarity 17.4%; Pred. No. 0.00067;
Matches 54; Conservative 131; Mismatches 124; Indels 2; Gaps 1;
QY 319 AAGTAAGCTGATTGGACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCA 378
Db 445 AATYAACRWAWANGAGAANKWKWAGSMRACARAGTTMAWGACWAMARRGWRCARGR--T 388
QY 379 GATGAATCTCCTAATCAAGATGGGAATGCCTCAGGTCAGGTAGCTAGCAAGAAAAACAAAG 438
Db 387 GSWGSKGGYRMWGRGMAAAKMRMAAGSYCGMTSYTSGSKMTGRKKGSMTKRKMPTY 328
QY 439 CAATTTACATAGATTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTG 498
Db 327 SGWWTSYKCTKTGKYTGWKSCKTRWTCTSWRKYMMWMSGCWARSMSKSWARSWSYMMAC 268
QY 499 GCTAACAAAAACAGAAAGAAAGAAACAGGAAAATGGAGGAGCGCTCTTGGACCTGATCT 558
Db 267 WCMWSASAYRARRSMYGARRSMRAGAGWWRARRGKRRARGKSSMMRSMRMSAGKA 208
QY 559 TGAAGACCTAAAACGCCAAGTACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGA 618
Db 207 RMCRRMWSCRRMSYSCMGSKCMSCRGTCAKWRYARYAKRYASSMGKYMGMGRWCYAKC 148
QY 619 ACAAGTCAGGG 629
Db 147 ARMVGYRSRS 137

RESULT 14

US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 655294
; GENERAL INFORMATION:

```
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

Query Match 2.1%; Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 0.7;
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
Y 342 AATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGAT 401
b 656477 AATTAGAAGAAAGAGAAGAAATTGAGGATATCAAGACTCAGATACAAAAT 656418
Y 402 GGGATGCCCTCAGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGG 461
b 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCCTAGGCAAGCAATTTATGTGG 656358
Y 462 ATCTCCAGAAATCAGAAACTGAAAGAGTTGATGACTGGCTAAACAAAAACAGA 513
b 656357 CAATTAAGCAGAAAAAACTAAAAAGCTTCTAAAGAGACGCGCTGGGAAATAGA 656306

ESULT 15
S-09-107-532A-1186
Sequence 1186, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
SEQUENCE CHARACTERISTICS:
LENGTH: 1179 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1179
SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match 2.1%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.0087;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 471 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAA 530
Db 602 AGCAAAGAATAAAGATTTTGATCGATCAGACAAAGAAAAATGGAGATACGATCGGAGGAA 661
QY 531 TGGAGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATA 590
Db 662 TTGTAGAAGTGCTCGTTGGAGGCGTTCCAGCTGGATTAGGAAGCTACGTACAATGGGACA 721
QY 591 AGGTGCTTCAAGAAGATCTAGAACAAGAACAACTCAGGGTCAATTCCTCACTCACATGG 650
Db 722 CGAAGCTAGATGCCAAAATCGCACAAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCGTAG 781
QY 651 TGGTGGTAGTTGATGAATCTAGTGG 675
Db 782 AATTGGGTCGGATTCACTTCTGG 806

Search completed: April 5, 2004, 08:24:46
Job time : 116.891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

1 nucleic - nucleic search, using sw model

in on: April 5, 2004, 04:53:30 ; Search time 509.455 Seconds
(without alignments)
1545.047 Million cell updates/sec

File: US-09-845-416-2_COPY_900_3000

Effect score: 2101

Sequence: 1 GAGCTATGCTACACAGG.....ACCTCAGCACTCTGGAAGAC 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Aligned: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match %	Length	DB ID	Description
1	2101	100.0	4182	10	US-09-845-416-2
2	2101	100.0	5149	10	US-09-845-416-27
3	1579.8	75.2	3999	10	US-09-845-416-6
4	1579.8	75.2	4966	10	US-09-845-416-28
5	1579.8	75.2	4990	10	US-09-845-416-34
6	1443	68.7	3858	10	US-09-845-416-9
7	1443	68.7	4825	10	US-09-845-416-29
8	1443	68.7	4848	10	US-09-845-416-35
9	1443	68.7	5060	10	US-09-845-416-36
10	1293	61.5	8689	15	US-10-149-736-42
11	1114	53.0	4414	10	US-09-845-416-32
12	1103.4	52.5	5417	15	US-10-149-736-39
13	1103.4	52.5	11443	15	US-10-149-736-44
14	1103.4	52.5	12057	15	US-10-149-736-47
15	1103.4	52.5	13957	9	US-09-782-378A-22

16	1103.4	52.5	13957	9	US-09-880-107-2284	Sequence 2284, Appl
17	1103.4	52.5	13957	15	US-10-149-736-1	Sequence 1, Appl
18	1103.4	52.5	14069	12	US-10-342-887-434	Sequence 434, Appl
19	1103.4	52.5	14082	12	US-10-342-887-981	Sequence 981, Appl
20	1103.4	52.5	14082	15	US-10-341-434-108	Sequence 108, Appl
21	1102	52.5	3446	10	US-09-845-416-14	Sequence 14, Appl
22	1100.2	52.4	11058	10	US-09-845-416-1	Sequence 1, Appl
23	1097.8	52.3	1991	10	US-09-845-416-3	Sequence 3, Appl
24	999	47.5	2169	10	US-09-845-416-4	Sequence 4, Appl
25	999	47.5	3531	10	US-09-845-416-10	Sequence 10, Appl
26	999	47.5	4498	10	US-09-845-416-30	Sequence 30, Appl
27	997	47.5	5339	15	US-10-149-736-40	Sequence 40, Appl
28	996	47.4	5462	15	US-10-149-736-41	Sequence 41, Appl
29	866.6	41.2	13815	15	US-10-149-736-2	Sequence 2, Appl
30	787	37.5	3510	10	US-09-845-416-12	Sequence 12, Appl
31	787	37.5	4476	10	US-09-845-416-31	Sequence 31, Appl
32	777	37.0	1667	10	US-09-845-416-7	Sequence 7, Appl
33	652	31.0	1821	10	US-09-845-416-13	Sequence 13, Appl
34	450	21.4	1340	10	US-09-845-416-11	Sequence 11, Appl
35	387	18.4	387	15	US-10-149-736-32	Sequence 32, Appl
36	361.4	17.2	10302	9	US-09-782-378A-23	Sequence 23, Appl
37	361.4	17.2	10302	15	US-10-149-736-3	Sequence 3, Appl
38	361.4	17.2	16531	14	US-10-101-510-667	Sequence 667, Appl
39	348	16.6	348	15	US-10-149-736-31	Sequence 31, Appl
40	336.6	16.0	11096	15	US-10-149-736-4	Sequence 4, Appl
41	333	15.8	333	15	US-10-149-736-9	Sequence 9, Appl
42	327	15.6	327	15	US-10-149-736-8	Sequence 8, Appl
43	322.4	15.3	333	15	US-10-149-736-10	Sequence 10, Appl
44	265	12.6	1434	10	US-09-845-416-15	Sequence 15, Appl
45	261	12.4	324	15	US-10-149-736-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match	100.0%	Score 2101;	DB 10;	Length 4182;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2101;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	GAGCTATGCTTACACACAGGCTGCTTATGTCCACCTCTGACCTACACGAGCCCAT	60	
Db	900	GAGCTATGCTTACACACAGGCTGCTTATGTCCACCTCTGACCTACACGAGCCCAT	959	
QY	61	TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG	120	
Db	960	TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG	1019	
QY	121	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	180	
Db	1020	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	1079	
QY	181	TGCTGAGGACACATTGCAAGCAAGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	240	

b 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1139
y 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 300
b 1140 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 1199
y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAATGAAGA 360
b 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAATGAAGA 1259
y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
b 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319
y 421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 480
b 1320 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 1379
y 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 540
b 1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 1439
y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAAATAGGTGCTTCA 600
c 1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAAATAGGTGCTTCA 1499
y 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
c 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 1559
y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGTATTGGG 720
c 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGTATTGGG 1619
y 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
c 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1679
y 781 CCTTCTCAAAATGSCAAACGCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTCAGA 840
b 1680 CCTTCTCAAAATGSCAAACGCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTCAGA 1739
y 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
b 1740 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1799
y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
b 1800 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1859
y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
b 1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1919
y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
b 1920 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1979
y 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
b 1980 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2039
y 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
b 2040 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC 2099
y 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 1260
b 2100 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 2159
y 1261 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACTGGATGAAACACAG 1320

Db 2160 AGACCTCCAAGGTGAAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAG 2219
Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 1380
Db 2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 2279
Qy 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2280 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2339
Qy 1441 TTTGGAAGCCAGTCTGACCCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2340 TTTGGAAGCCAGTCTGACCCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2399
Qy 1501 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 2400 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 2459
Qy 1561 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 2460 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2519
Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 1680
Db 2520 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 2579
Qy 1681 AGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 1740
Db 2580 AGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 2639
Qy 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAA 1800
Db 2640 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAA 2699
Qy 1801 CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT 1860
Db 2700 CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT 2759
Qy 1861 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
Db 2760 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 2819
Qy 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 2820 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 2879
Qy 1981 GGCACCTTCGAGGAGAAAATTGCGCCTCTGAAAAGAGAACCGTGAGCCACGTCATGACCTTGC 2040
Db 2880 GGCACCTTCGAGGAGAAAATTGCGCCTCTGAAAAGAGAACCGTGAGCCACGTCATGACCTTGC 2939
Qy 2041 TCGCCAGCTTACCATTGCGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAGA 2100
Db 2940 TCGCCAGCTTACCATTGCGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAGA 2999
Qy 2101 C 2101
Db 3000 C 3000

RESULT 2
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; TITLE OF INVENTION: DE1142
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 27									
LENGTH: 5149									
TYPE: DNA									
ORGANISM: Homo sapiens									
S-09-845-416-27									
Y	1	GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCA	60	Query Match					
b	1657	GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCA	1716	Best Local Similarity					
Y	61	TCCTTTCACAGCAATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTTGATGGAGAG	120	Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
b	1717	TCCTTTCACAGCAATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTTGATGGAGAG	1776						
Y	121	TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC	180						
b	1777	TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC	1836						
Y	181	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240						
b	1837	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1896						
Y	241	CCAGTTTTCATCTACATGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGGTGG	300						
b	1897	CCAGTTTTCATCTACATGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGGTGG	1956						
Y	301	TAATATTCTACAAATTTGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA	360						
b	1957	TAATATTCTACAAATTTGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA	2016						
Y	361	AACTGAAGTACAAAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	420						
b	2017	AACTGAAGTACAAAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	2076						
Y	421	TAGCATGGAAACAAAGCAATTTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT	480						
b	2077	TAGCATGGAAACAAAGCAATTTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT	2136						
Y	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAGTGAAGA	540						
b	2137	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAGTGAAGA	2196						
Y	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAAGTGCCTTCA	600						
b	2197	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAAGTGCCTTCA	2256						
Y	601	AGAAGATCTAGAAACAAAGCAAGTCAAGGTCAATTTCTCACTCACATGGTGGGTAGT	660						
b	2257	AGAAGATCTAGAAACAAAGCAAGTCAAGGTCAATTTCTCACTCACATGGTGGGTAGT	2316						
Y	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAAACAACTTAAGGTATTGGG	720						
b	2317	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAAACAACTTAAGGTATTGGG	2376						
Y	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTCTTTTACAAAGACAT	780						
b	2377	AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTCTTTTACAAAGACAT	2436						
Y	781	CCTTCTCAAATGGCAACGCTCTTACTGAAGAAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA	840						
b	2437	CCTTCTCAAATGGCAACGCTCTTACTGAAGAAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA	2496						
Y	841	AAAAGAAGATGCAAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	900						
b	2497	AAAAGAAGATGCAAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	2556						
Y	901	ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT	960						

Db	2557	ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT	2616
QY	961	GGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTCAACACCTGAAGAATAAGTCAGTGAC	1020
Db	2617	GGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTCAACACCTGAAGAATAAGTCAGTGAC	2676
QY	1021	CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTAGTCCAAA	1080
Db	2677	CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTAGTCCAAA	2736
QY	1081	ACTTGAAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCTCGACCTGGA	1140
Db	2737	ACTTGAAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCTCGACCTGGA	2796
QY	1141	AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC	1200
Db	2797	AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC	2856
QY	1201	CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA	1260
Db	2857	CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA	2916
QY	1261	AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG	1320
Db	2917	AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG	2976
QY	1321	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT	1380
Db	2977	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT	3036
QY	1381	GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA	1440
Db	3037	GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA	3096
QY	1441	TTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT	1500
Db	3097	TTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT	3156
QY	1501	GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC	1560
Db	3157	GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC	3216
QY	1561	AGCAGTTTCAAGACGACATGTACATAGGSCCTTCAAGAGGGAATTTGAAAACTAAAGA	1620
Db	3217	AGCAGTTTCAAGACGACATGTACATAGGSCCTTCAAGAGGGAATTTGAAAACTAAAGA	3276
QY	1621	ACCTGTAATCATGATGACTCTTGAGACTGTACCAATATTCTGACAGAGCAGCCTTTTGA	1680
Db	3277	ACCTGTAATCATGATGACTCTTGAGACTGTACCAATATTCTGACAGAGCAGCCTTTTGA	3336
QY	1681	AGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAA	1740
Db	3337	AGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAA	3396
QY	1741	TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA	1800
Db	3397	TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA	3456
QY	1801	CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCCAGGA	1860
Db	3457	CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCCAGGA	3516
QY	1861	TCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC	1920
Db	3517	TCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC	3576
QY	1921	CTGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	1980
Db	3577	CTGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	3636
QY	1981	GGCACTTCGAGGAGAAATTTGGCCTCTCTGAAAGAGAAACGTCAGCCACGTCATGACCTTGC	2040
Db	3637	GGCACTTCGAGGAGAAATTTGGCCTCTCTGAAAGAGAAACGTCAGCCACGTCATGACCTTGC	3696

Y 2041 TCGCCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATAAACCCTCAGCACCTCTGGAAGA 2100
b 3697 TCGCCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATAAACCCTCAGCACCTCTGGAAGA 3756
Y 2101 C 2101
b 3757 C 3757

ESULT 3
S-09-845-416-6
Sequence 6, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO

TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-6

Query Match 75.2%; Score 1579.8; DB 10; Length 3999;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 60
b 900 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 959
Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
b 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1019
Y 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGTATATCGTGGCTTCTTC 180
b 1020 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGTATATCGTGGCTTCTTC 1079
Y 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGGAAAGTGGTGAAGA 240
b 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGGAAAGTGGTGAAGA 1139
Y 241 CCAGTTTCATACCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
b 1140 CCAGTTTCATACCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1199
Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
b 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
b 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1319
Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
b 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
Y 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAGAAACAAGGAAAATGGAGGAAGA 540
b 1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAGAAACAAGGAAAATGGAGGAAGA 1439
Y 541 GCCTCTGGACCTGATCTTGAAGACCTTAAACGGCAAGTACAACAATAGGTGCTTCA 600

Db 1440 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTACAACACATAAAGTGCTTCA 1499
QY 601 AGAAGATCTAGAAACAACAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 660
Db 1500 AGAAGATCTAGAAACAACAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACCTAAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACCTAAAGGTATTGGG 1619
QY 721 AGATCGATGGCAACAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1620 AGATCGATGGCAACAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACCA 1679
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 1680 GCCTGAC-----CTAGCTCCTGGACTGACCATAATTGGAGCCTCTCCTACTCA 1727
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAAATGAATGTT 900
Db 1728 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAAACCTGCCATCT- 1777
QY 901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1778 ----- 1777
QY 961 GGGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTTCAACACACTGAAGAATAAGTCAGTGAC 1020
Db 1778 -----CAAAACCTAGAAATGCCATCTTCTTGTATGTTGGAG----- 1812
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGGCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 1813 ----- 1812
QY 1081 ACTTGAAGAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGATTCGCCCTGGACCTGGA 1140
Db 1813 -----GTACCTACTCATAGATTACTGCAACAGATTCGCCCTGGACCTGGA 1856
QY 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
Db 1857 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1916
QY 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAAACAATGGCA 1260
Db 1917 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAAACAATGGCA 1976
QY 1261 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
Db 1977 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2036
QY 1321 CCARAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCA 1380
Db 2037 CCARAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCAGTCTCTTACAAAAGACGTTT 2096
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2097 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 2156
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2157 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGT 2216
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCC 1560
Db 2217 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCC 2276
QY 1561 AGCAGTTCAGAAGCAGAACCATGTACATAGGGCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
Db 2277 AGCAGTTCAGAAGCAGAACCATGTACATAGGGCTTCAAGAGGGGAATTGAAAACTAAAGA 2336
QY 1621 ACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 2337 ACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 2396

1681 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTGTAGGAGAGAGCCAGAA 1740
|||||
2397 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTGTAGGAGAGAGCCAGAA 2456
|||||
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA 1800
|||||
2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA 2516
|||||
1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 1860
|||||
2517 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 2576
|||||
1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
|||||
2577 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 2636
|||||
1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 1980
|||||
2637 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 2696
|||||
1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 2040
|||||
2697 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 2756
|||||
2041 TCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 2100
|||||
2757 TCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 2816
|||||
2101 C 2101
2817 C 2817

ESULT 4
S-09-845-416-28
Sequence 28, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 28
LENGTH: 4966
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-28

Query Match 75.2%; Score 1579.8; DB 10; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
|||||
1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1716
|||||
61 TCCTTCACAGCATTTGGAAGCTCCTGGAAGACAAGTCAATTGGCAGTTTCATTGATGGAGAG 120
|||||
1717 TCCTTCACAGCATTTGGAAGCTCCTGGAAGACAAGTCAATTGGCAGTTTCATTGATGGAGAG 1776
|||||
121 TGAAGTAAACCTTGACCGGTTATCAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
|||||
1777 TGAAGTAAACCTTGACCGGTTATCAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1836
|||||
181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||

Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
Db 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
|||||
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
Db 2077 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAAATGGAGGAAGA 540
|||||
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAAATGGAGGAAGA 2196
QY 541 GCGTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
|||||
Db 2197 GCGTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACRAAGAACAAAGTCAGGGTCAATCTCTCACTCAGATGGTGGTAGT 660
|||||
Db 2257 AGAAGATCTAGAACRAAGAACAAAGTCAGGGTCAATCTCTCACTCAGATGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
|||||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 780
|||||
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACCA 2436
QY 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA 840
|||||
Db 2437 GCGTGAC-----CTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCCTACTCA 2484
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
|||||
Db 2485 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2534
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 2535 ----- 2534
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACACTGAAGAATAAGTCAGTGAC 1020
|||||
Db 2535 -----CCAAACTAGAAATGCCATCTTCTCCTGATGTGGAG----- 2569
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1080
Db 2570 ----- 2569
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
|||||
Db 2570 -----GTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2613
QY 1141 AAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
|||||
Db 2614 AAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGCTAC 2673
QY 1201 CCCTAAGGAAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGTGAATGAAAAATGGCA 1260
|||||
Db 2674 CCCTAAGGAAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGTGAATGAAAAATGGCA 2733
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTCTGGATGAAAAACAG 1320
|||||
Db 2734 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTCTGGATGAAAAACAG 2793
|||||

```
1321 CCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTCTCTGTACAAAAGACGTTT 1380
1322 |||||||
2794 CCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTCTCTGTACAAAAGACGTTT 2853
2795 |||||||
1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
1382 |||||||
2854 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 2913
2855 |||||||
1441 TTTGGAAGCCAGTTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 1500
1442 |||||||
2914 TTTGGAAGCCAGTTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 2973
2915 |||||||
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 1560
1502 |||||||
2974 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 3033
2975 |||||||
1561 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGGAATTGAAAACCTAAAGA 1620
1562 |||||||
3034 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGGAATTGAAAACCTAAAGA 3093
3035 |||||||
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
1622 |||||||
3094 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3153
3095 |||||||
1681 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
1682 |||||||
3154 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3213
3155 |||||||
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
1742 |||||||
3214 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3273
3215 |||||||
1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAATCCAGGAAT 1860
1802 |||||||
3274 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAATCCAGGAAT 3333
3275 |||||||
1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGTGAATCAAGGGATC 1920
1862 |||||||
3334 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGTGAATCAAGGGATC 3393
3335 |||||||
1921 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
1922 |||||||
3394 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3453
3395 |||||||
1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGC 2040
1982 |||||||
3454 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGC 3513
3455 |||||||
2041 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGA 2100
2042 |||||||
3514 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGA 3573
3515 |||||||
2101 C 2101
3574 C 3574
```

RESULT 5

```
JS-09-845-416-34
: Sequence 34, Application US/09845416
: Publication No. US20030171312A1
: GENERAL INFORMATION:
: APPLICANT: XIAO, XIAO
: TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
: FILE REFERENCE: DB1142
: CURRENT APPLICATION NUMBER: US/09/845,416
: CURRENT FILING DATE: 2001-04-30
: PRIOR APPLICATION NUMBER: 60/200,777
: PRIOR FILING DATE: 2000-04-28
: NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
```

```
: SEQ ID NO 34
: LENGTH: 4990
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-845-416-34

Query Match          75.2%; Score 1579.8; DB 10; Length 4990;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGCCCAT 60
Db |||||||
1681 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGCCCAT 1740
1682 |||||||

QY 61 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db |||||||
1741 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1800
1742 |||||||

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db |||||||
1801 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAAGAAGAAGTATTATCGTGGCTTCTTTC 1860
1802 |||||||

QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Db |||||||
1861 TGCTGAGGACACATTGCAAGCAAGGAGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1920
1862 |||||||

QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 300
Db |||||||
1921 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 1980
1922 |||||||

QY 301 TAATATTCTACAATTTGGGAATTAAGCTGATTGGAAACAGGAAAAATTTACAGAAGATGAAGA 360
Db |||||||
1981 TAATATTCTACAATTTGGGAATTAAGCTGATTGGAAACAGGAAAAATTTACAGAAGATGAAGA 2040
1982 |||||||

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTAAGATGGGAATGCCTCAGGGTAGC 420
Db |||||||
2041 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTAAGATGGGAATGCCTCAGGGTAGC 2100
2042 |||||||

QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db |||||||
2101 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2160
2102 |||||||

QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 540
Db |||||||
2161 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 2220
2162 |||||||

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACTAAGGTGCTTCA 600
Db |||||||
2221 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACTAAGGTGCTTCA 2280
2222 |||||||

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db |||||||
2281 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 2340
2282 |||||||

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 720
Db |||||||
2341 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 2400
2342 |||||||

QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 780
Db |||||||
2401 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACCA 2460
2402 |||||||

QY 781 CCTTCTCAAATGGCAACGTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 840
Db |||||||
2461 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2508
2462 |||||||

QY 841 AAAAGAGATGCAGTGAACAGATTTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
Db |||||||
2509 GACTGTACTCTGCTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2558
2510 |||||||

QY 901 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db |||||||
2559 ----- 2558
```


961 GGGCAAACTGTATTCACTCRAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
2559 -----CCAAACTAGAAATGCCATCTTCTTGATGTTGGAG----- 2593
1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTAGTCCAAA 1080
2594 ----- 2593
1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
2594 -----GTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2637
1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTAC 1200
2638 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTAC 2697
1201 CCGTAAGGAAAGGCTCCTAGAAAGTCCAAAGGAGTAAAGAGCTGATGAACAATGGCA 1260
2698 CCGTAAGGAAAGGCTCCTAGAAAGTCCAAAGGAGTAAAGAGCTGATGAACAATGGCA 2757
1261 AGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
2758 AGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2817
1321 CCAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCGAGTCCCTGTTACAAGACGTTT 1380
2818 CCAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCGAGTCCCTGTTACAAGACGTTT 2877
1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAGTCTCTCAACATTAGGTCCCA 1440
2878 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAGTCTCTCAACATTAGGTCCCA 2937
1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
2938 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2997
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCC 1560
2998 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCC 3057
1561 AGCAGTTCAGAAACGAGATGTACATAGGGCTTCAAGAGGGAATTTGAAACTAAAGA 1620
3058 AGCAGTTCAGAAACGAGATGTACATAGGGCTTCAAGAGGGAATTTGAAACTAAAGA 3117
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTGGA 1680
3118 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTGGA 3177
1681 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGGCCAGAA 1740
3178 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGGCCAGAA 3237
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTTGAA 1800
3238 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTTGAA 3297
1801 CCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 1860
3298 CCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 3357
1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 1920
3358 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 3417
1921 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
3418 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3477
1981 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040
3478 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3537

2041 TCGCCAGCTTACCACATTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCACTCTGGGAGA 2100
3538 TCGCCAGCTTACCACATTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCACTCTGGGAGA 3597
2101 C 2101
3598 C 3598
RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
Query Match 68.7%; Score 1443; DB 10; Length 3858;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACACGGAGCCCAAT 60
Db 900 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACACGGAGCCCAAT 959
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTAATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTAATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTCGAAGCACCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTCGAAGCACCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 300
Db 1140 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAGCAATTTTACATAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAGCAATTTTACATAGTTTTTAATGGATCTCCAGAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 1439
QY 541 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAAGGTGCTTCA 600
Db 1440 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAAGGTGCTTCA 1499

Y 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 660
b 1500 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1559
Y 661 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTTGGAAAGCAAACTTAAGGTATTGGG 720
b 1560 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTTGGAAAGCAAACTTAAGGTATTGGG 1619
Y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAGACAT 780
b 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAGAC-- 1677
Y 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTGTGTCATGGCTTTCAGA 840
b 1678 ----- 1677
Y 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 900
b 1678 ----- 1677
Y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
b 1678 ----- 1677
Y 961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
b 1678 ----- 1677
Y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
b 1678 ----- 1677
Y 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
b 1678 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1715
Y 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCCTACAGGATGCTAC 1200
b 1716 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCCTACAGGATGCTAC 1775
Y 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
b 1776 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 1835
Y 1261 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320
b 1836 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1895
Y 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 1380
b 1896 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 1955
Y 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
b 1956 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA 2015
Y 1441 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
b 2016 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2075
Y 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 1560
b 2076 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 2135
Y 1561 AGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
b 2136 AGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGA 2195
Y 1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
b 2196 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 2255

QY 1681 AGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1740
Db 2256 AGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 2315
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGAAAAATTGAA 1800
Db 2316 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGAAAAATTGAA 2375
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 1860
Db 2376 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 2435
QY 1861 TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2436 TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 2495
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 2496 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 2555
QY 1981 GGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
Db 2556 GGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2615
QY 2041 TCGCAGCTTACCACTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
Db 2616 TCGCAGCTTACCACTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2675
QY 2101 C 2101
Db 2676 C 2676

RESULT 7
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 68.7%; Score 1443; DB 10; Length 4825;
Best local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGAGGCCATT 60
Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGAGGCCATT 1716
QY 61 TCCTTCACAGCATTTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1896

Y 241 CCAGTTTCACTACTGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
b 1897 CCAGTTTCACTACTGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1956
Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
b 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 2016
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
b 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
Y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 480
b 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 2136
Y 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 540
b 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 2196
Y 541 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 600
b 2197 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 2256
Y 601 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 660
b 2257 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 2316
Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAGAACAACATTAAGTATTTGGG 720
b 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAGAACAACATTAAGTATTTGGG 2376
Y 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
b 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGAC-- 2434
Y 781 CCTTCTCAAATGSCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGSCCTTCAGA 840
b 2435 ----- 2434
Y 841 AAAAGAAGATGCGATGAACAAGATTCACACAACCTGGCTTTAAAGATCAAATGAAATGTT 900
b 2435 ----- 2434
Y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
b 2435 ----- 2434
Y 961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAAACACTGAAGAATAAGTCAGTGAC 1020
b 2435 ----- 2434
Y 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
b 2435 ----- 2434
Y 1081 ACTTGAAAAGATPACAGCACAGACTCATAGATTACTGCAACAGTTCCCTTGGACCTGGA 1140
b 2435 -----ACTCATAGATTACTGCAACAGTTCCCTTGGACCTGGA 2472
Y 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 1200
b 2473 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 2532
Y 1201 CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 1260
b 2533 CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 2592
Y 1261 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
b 2593 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2652

QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTT 1380
Db 2653 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTT 2712
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2713 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2772
QY 1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2773 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2832
QY 1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCAGCTTTCC 1560
Db 2833 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCAGCTTTCC 2892
QY 1561 AGCAGTTTCAAGAGCGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
Db 2893 AGCAGTTTCAAGAGCGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 2952
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 2953 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3012
QY 1681 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGAGAGAGAGAG 1740
Db 3013 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGAGAGAGAGAG 3072
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAA 1800
Db 3073 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAA 3132
QY 1801 CCTGCACCTCGGCTGAGTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860
Db 3133 CCTGCACCTCGGCTGAGTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 3192
QY 1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
Db 3193 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3252
QY 1921 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 3253 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3312
QY 1981 GGCACCTTCGAGGAGAAAAATTGCGCTCTGAAAAGAGAACGTCGACCGTCAATGACCTTGC 2040
Db 3313 GGCACCTTCGAGGAGAAAAATTGCGCTCTGAAAAGAGAACGTCGACCGTCAATGACCTTGC 3372
QY 2041 TCGCCAGCTTACCACTTTGGGCATTTGAGTCTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGA 2100
Db 3373 TCGCCAGCTTACCACTTTGGGCATTTGAGTCTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGA 3432
QY 2101 C 2101
Db 3433 C 3433

RESULT 8
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35

LENGTH: 4848									
TYPE: DNA									
ORGANISM: Homo sapiens									
S-09-845-416-35									
Query Match		68.7%;	Score 1443;	DB 10;	Length 4848;				
Best Local Similarity		84.6%;	Pred. No. 0;						
Matches 1777;		Conservative 0;	Mismatches 0;	Indels 324;	Gaps 1;				
y	1	GAGCTATGCCTACACACAGGCTGCTTATGTCA	CACCACCTCTGACCCCTACACGGAGCCCAT	60					
b	1680	GAGCTATGCCTACACACAGGCTGCTTATGTCA	CACCACCTCTGACCCCTACACGGAGCCCAT	1739					
y	61	TCCTTCACAGCATTTGGRAAGCTCCTGAAGACA	AGTCATTTGGCAGTTTCATTTGATGGAGAG	120					
b	1740	TCCTTCACAGCATTTGGRAAGCTCCTGAAGACA	AGTCATTTGGCAGTTTCATTTGATGGAGAG	1799					
y	121	TGAAGTAAACCTGGACCGTTATCAAACAGCTTT	TAGAAGAGTATTATCGTGGCTTCTTTC	180					
b	1800	TGAAGTAAACCTGGACCGTTATCAAACAGCTTT	TAGAAGAGTATTATCGTGGCTTCTTTC	1859					
y	181	TGCTGAGGACACATTTGCAAGCACAAGGAGAGA	TTTTCTAATGATGTGGAAGTGGTGAAGA	240					
b	1860	TGCTGAGGACACATTTGCAAGCACAAGGAGAGA	TTTTCTAATGATGTGGAAGTGGTGAAGA	1919					
y	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGAT	TTTGACAGCCCATCAGGGCCGGGTTGG	300					
b	1920	CCAGTTTTCATCTCATGAGGGGTACATGATGGAT	TTTGACAGCCCATCAGGGCCGGGTTGG	1979					
y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGA	ACAGGAAAAATTATCAGAAGATGAAGA	360					
b	1980	TAATATTCTACAATTGGGAAGTAAGCTGATTGGA	ACAGGAAAAATTATCAGAAGATGAAGA	2039					
y	361	AACTGAAGTACAAGACAGATGAATCTCCTAAAT	TCAAGATGGGAATGCCTCAGGGTAGC	420					
b	2040	AACTGAAGTACAAGACAGATGAATCTCCTAAAT	TCAAGATGGGAATGCCTCAGGGTAGC	2099					
y	421	TAGCATGGAAAAACAAGCAATTTACATAGAGTTT	TAAATGGATCTCCAGAATCAGAAAACT	480					
b	2100	TAGCATGGAAAAACAAGCAATTTACATAGAGTTT	TAAATGGATCTCCAGAATCAGAAAACT	2159					
y	481	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGA	AGAAAGAAACAAGGAAAAATGGAGGAAGA	540					
b	2160	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGA	AGAAAGAAACAAGGAAAAATGGAGGAAGA	2219					
y	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGC	CAAGTACAAACAACATAAGGTGCTTCA	600					
b	2220	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGC	CAAGTACAAACAACATAAGGTGCTTCA	2279					
y	601	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCA	ATTCTCTCACTCACATGGTGGTGTAGT	660					
b	2280	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCA	ATTCTCTCACTCACATGGTGGTGTAGT	2339					
y	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTT	TGGBAAGAACAACTTAAGGTATTGGG	720					
b	2340	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTT	TGGBAAGAACAACTTAAGGTATTGGG	2399					
y	721	AGATCGATGGGCAACATCTGTAGATGGACAGAGA	CCGCTGGGTTCTTTTACAGACAT	780					
b	2400	AGATCGATGGGCAACATCTGTAGATGGACAGAGA	CCGCTGGGTTCTTTTACAGAC--	2457					
y	781	CCTTCTCAAATGGCAACGCTCTTACTGAAGAAACA	GTCCTTTTATGTGCATGGCTTTCAGA	840					
b	2458	-----	-----	2457					
y	841	AAAAGAAGATGCAGTGAACAAGATTTCACACAAC	TGGCTTTTAAAGATCAAAATGAAATGTT	900					
b	2458	-----	-----	2457					
y	901	ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGA	TCTAGAAAAAGAAAAAGCAATCCAT	960					
b	2458	-----	-----	2457					

QY	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1020
Db	2458	-----	2457
QY	1021	CCAGAAAGACGGAAGCATGGCTGGATAAATTTGCCCCGGTGTGGGATAATTTAGTCCAAAA	1080
Db	2458	-----	2457
QY	1081	ACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA	1140
Db	2458	-----ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA	2495
QY	1141	AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC	1200
Db	2496	AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC	2555
QY	1201	CCGTAAGGAAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA	1260
Db	2556	CCGTAAGGAAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA	2615
QY	1261	AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTTGGATGAAAACAG	1320
Db	2616	AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTTGGATGAAAACAG	2675
QY	1321	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAAGACGTTT	1380
Db	2676	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAAGACGTTT	2735
QY	1381	GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA	1440
Db	2736	GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA	2795
QY	1441	TTTGGAAAGCCAGTTCTGACCAGTGGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT	1500
Db	2796	TTTGGAAAGCCAGTTCTGACCAGTGGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT	2855
QY	1501	GTGGCTACAGCTGAAAGATGATGAATTTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC	1560
Db	2856	GTGGCTACAGCTGAAAGATGATGAATTTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC	2915
QY	1561	AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAGA	1620
Db	2916	AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAGA	2975
QY	1621	ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTTCTGACAGAGCAGCCCTTTGGA	1680
Db	2976	ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTTCTGACAGAGCAGCCCTTTGGA	3035
QY	1681	AGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGAGAGAGAGAG	1740
Db	3036	AGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGAGAGAGAGAG	3095
QY	1741	TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA	1800
Db	3096	TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA	3155
QY	1801	CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT	1860
Db	3156	CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT	3215
QY	1861	TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC	1920
Db	3216	TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC	3275
QY	1921	CTGGCAGCCCGTGGCGATCTCTCTATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	1980
Db	3276	CTGGCAGCCCGTGGCGATCTCTCTATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	3335
QY	1981	GGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAAAGTGAGCCACGTCATGACCTTGC	2040
Db	3336	GGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAAAGTGAGCCACGTCATGACCTTGC	3395
QY	2041	TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATATACCTCAGCACTCTTGGAAGA	2100

||||| 3396 TCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 3455
2101 C 2101
3456 C 3456

RESULT 9
3-09-845-416-36
Sequence 36, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 5060
TYPE: DNA
ORGANISM: Homo sapiens
3-09-845-416-36

Query Match 68.7%; Score 1443; DB 10; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGAGCCCAT 60
1892 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGAGCCCAT 1951

61 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTGAGAG 120
1952 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTGAGAG 2011

121 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
2012 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 2071

181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
2072 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131

241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTTGG 300
2132 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTTGG 2191

301 TAATATTCTACAATTGGGAAGTAACTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 360
2192 TAATATTCTACAATTGGGAAGTAACTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2251

361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
2252 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2311

421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
2312 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2371

481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 540
2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 2431

541 GCCTCTTGGACCTGATCTTTGAAGACCTAAAACGCCAAGTACAACAACATAGGTGCTTCA 600
2432 GCCTCTTGGACCTGATCTTTGAAGACCTAAAACGCCAAGTACAACAACATAGGTGCTTCA 2491

Qy 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTGTAGT 660
Db 2492 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTGTAGT 2551

Qy 661 TGATGAATCTAGTCGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2552 TGATGAATCTAGTCGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2611

Qy 721 AGATCGATGGGCAAAACATCTGTAGATGAGACAGACCGCTGCTTTTACAAAGACAT 780
Db 2612 AGATCGATGGGCAAAACATCTGTAGATGAGACAGACCGCTGCTTTTACAAAGAC-- 2669

Qy 781 CCTTCTCAAATGSCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTTCAGA 840
Db 2670 ----- 2669

Qy 841 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 2670 ----- 2669

Qy 901 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db 2670 ----- 2669

Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2670 ----- 2669

Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCGGCTGCTGGGATAATTTAGTCCAAA 1080
Db 2670 ----- 2669

Qy 1081 ACTTGAAAAAGAGTACAGACACAGACTCATAGATTACTGCAACAGTTCCCTGACCTGGA 1140
Db 2670 -----ACTCATAGATTACTGCAACAGTTCCCTGACCTGGAACCTGGA 2707

Qy 1141 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
Db 2708 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2767

Qy 1201 CCGTAAGGAAAGGCTCTAGAAGACTCCAAGGGAGTAAAGAGTGAATAAGCAATGGCA 1260
Db 2768 CCGTAAGGAAAGGCTCTAGAAGACTCCAAGGGAGTAAAGAGTGAATAAGCAATGGCA 2827

Qy 1261 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACTGGATGAAACAG 1320
Db 2828 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACTGGATGAAACAG 2887

Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTT 1380
Db 2888 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTT 2947

Qy 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2948 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3007

Qy 1441 TTTGGAAGCCAGTTCTGACCCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGT 1500
Db 3008 TTTGGAAGCCAGTTCTGACCCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGT 3067

Qy 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGCGGACTTCC 1560
Db 3068 GTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGCGGACTTCC 3127

Qy 1561 AGCAGTTCAGAACGAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 3128 AGCAGTTCAGAACGAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 3187

Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTTGACAGAGCAGCCTTTGGA 1680
Db 3188 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTTGACAGAGCAGCCTTTGGA 3247

Qy 1681 AGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGAGAGAGAGAG 1740

b 3248 AGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCTCGAGGAGAGAGCCAGAA 3307
Y 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
b 3308 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3367
Y 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
b 3368 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3427
Y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
b 3428 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3487
Y 1921 CTGGCAGCCCGTGGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
b 3488 CTGGCAGCCCGTGGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3547
Y 1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 2040
b 3548 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 3607
Y 2041 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATATAACCTCAGCACCTCTGGAAGA 2100
b 3608 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATATAACCTCAGCACCTCTGGAAGA 3667
Y 2101 C 2101
b 3668 C 3668

ESULT 10
S-10-149-736-42
Sequence 42, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42
LENGTH: 8689
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
S-10-149-736-42

Query Match 61.5%; Score 1293; DB 15; Length 8689;
Best Local Similarity 72.5%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;
Y 1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCTACACGGAGCCCAATT 60
b 1099 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCTACACGGAGCCCAATT 1158
Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
Y 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGACATTTCTAATGATGTGGAAGTGGTGAAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGACATTTCTAATGATGTGGAAGTGGTGAAAGA 1338
QY 241 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTCGGAACAGGAAAAATATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTCGGAACAGGAAAAATATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAGAACAAAGTCAGGTCAATTTCTCTACTCATCGTGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAGAACAAAGTCAGGTCAATTTCTCTACTCATCGTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTCA 840
Db 1879 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTCA 1938
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCTCAAAACAAGATCTTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCTCAAAACAAGATCTTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAG----- 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTTCACAGCAGCCTGACCTAGCTCCTGGACTGACCAC 2238
QY 1103 ----- 1102
Db 2239 TATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACACACTGTGGTTACTAAGGA 2298
QY 1103 ----- 1102
Db 2299 AACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGTGTTGGAGGTACCTGCTCTGTC 2358
QY 1103 ----- 1102

2359 AGATTTC AACCGGCTTGGACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTAT 2418
1103 -----
2419 AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478
1103 -----
2479 GAAGGCAACAATGCAGGATTTGGAACAGAGCGCTCCCGAGTTGGAAGAACTCATTAACCGC 2538
1103 -----
2539 TGCCCAAAATTTGAAAACAAGACCAGCAATCAAGAGGCTAGAACAAATCATTAACGGATCG 2598
1103 -----
2599 AATTGAAAGAATTCAGAAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCA 2658
1103 -----
2659 ACAGTTGAATGAAATGTTAAAGGATTCAACACAAATGGCTGGAAGCTAAGGAAGAAGCTGA 2718
1103 -----
2719 GCAGGTCTTAGGACAGGCCAGAGCCAAAGCTTGAGTCATGGAAGGAGGGTCCCTATACAGT 2778
1103 -----
2779 AGATGCAATCCAAAAGAAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG 2838
1103 -----
2839 GCAGACAAATGTAGATGTGGCAAAATGACTTGGCCCTGAAAATTCTCCGGGATTATTCTGC 2898
1103 -----
2899 AGATGATACCAGAAAAGTCCACATGATAACAGAGAATATCAATGCTCTTGGAGAAAGCAT 2958
1103 -----
2959 TCATAAAAGGGTGAGTGAGCGAGAGGCTGCTTTTGGAAAGAACTCATAGATTACTGCAACA 3018
1123 GTTCCCTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACCTGCCAA 1182
3019 GTTCCCTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACCTGCCAA 3078
1183 TGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGA 1242
3079 TGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGA 3138
1243 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCA 1302
3139 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCA 3198
1303 CAACCTGGATGAAAACAGCCAAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1362
3199 CAACCTGGATGAAAACAGCCAAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 3258
1363 CCTGTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAGTC 1422
3259 CCTGTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAGTC 3318
1423 TCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTC 1482
3319 TCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTC 3378
1483 TCTCAGGAACTTCTGTTGGCTACAGCTGAAGATGATGAATTAAGCCCGCAGGACCC 1542
3379 TCTCAGGAACTTCTGTTGGCTACAGCTGAAGATGATGAATTAAGCCCGCAGGACCC 3438
1543 TATTGAGGCGACTTTCACGACGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAACAG 1602

Db 3439 TATTGGAGCGACTTTCCAGCAGTTTCAGAAACGATGTACATAGGGCCTTCAAGAG 3498
Qy 1603 GGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCT 1662
Db 3499 GGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCT 3558
Qy 1663 GACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAAGAGCCCAGAGAGCTGCCTCC 1722
Db 3559 GACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAAGAGCCCAGAGAGCTGCCTCC 3618
Qy 1723 TGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 1782
Db 3619 TGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 3678
Qy 1783 TGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 1842
Db 3679 TGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 3738
Qy 1843 TGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGC 1902
Db 3739 TGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGC 3798
Qy 1903 TGAGGTGATCAAGGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCCAAGA 1962
Db 3799 TGAGGTGATCAAGGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCCAAGA 3858
Qy 1963 TCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCTCTGAAAGAGAACCTGAG 2022
Db 3859 TCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCTCTGAAAGAGAACCTGAG 3918
Qy 2023 CCACGTCAATGACCTTGTCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACGTCAATGACCTTGTCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAA 3978
Qy 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 11

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 53.0%; Score 1114; DB 10; Length 4414;
Best Local Similarity 97.0%; Pred. No. 2.9e-313;
Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGAGGCCATT 60
Db 1657 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGAGGCCATT 1716
Qy 61 TCCTTACAGCATTGGAAGCTCTGAAGACAAGTCATTGTCAGTTGGCAGTTTCATTGATGGAGAG 120
Db 1717 TCCTTACAGCATTGGAAGCTCTGAAGACAAGTCATTGTCAGTTGGCAGTTTCATTGATGGAGAG 1776
Qy 121 TGAAGTAAACCTGGACCGCTTATCAACACAGCTTTAGAGAAGTATTATTCGTGGCTTCTTTC 180

b 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1836
Y 181 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 240
b 1837 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 1896
Y 241 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 300
b 1897 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 1956
Y 301 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
b 1957 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
Y 361 AACTGAAGTACAAGACGAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
b 2017 AACTGAAGTACAAGACGAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
Y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 480
b 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 2136
Y 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 540
b 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 2196
Y 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAAGGTGCTTCA 600
b 2197 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAAGGTGCTTCA 2256
Y 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
b 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 2316
Y 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
b 2317 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
Y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTTCCTTTACAGACAT 780
b 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTTCCTTTACAGACAT 2436
Y 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTTCA 840
b 2437 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTTCA 2496
Y 841 AAAAGAAGATGCAGTGAAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 900
b 2497 AAAAGAAGATGCAGTGAAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 2556
Y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
b 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 2616
Y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 1020
b 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 2676
Y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAA 1080
b 2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAA 2736
Y 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCTTGGACCTGGA 1140
b 2737 ACTTGAAAAGAGTACAGCACAGACCCTTGAAAGACTCCAGAACTTCAAGAGGCCACGGA 2796
Y 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGA 1170
b 2797 TGAGCTGGACCTCAAGCTGCGCCAAGCTGA 2826

ESULT 12

US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match 52.5%; Score 1103.4; DB 15; Length 5417;
Best Local Similarity 99.9%; Pred. No. 4e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCATTT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCATTT 1158
QY 61 TCCTTACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAACACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1398
QY 301 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAAGGTGCTTCA 600
Db 1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720

b 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1818
y 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
b 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
y 781 CCTTCTCAAAATGSCAAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 840
b 1879 CCTTCTCAAAATGSCAAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 1938
y 841 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
b 1939 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
b 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
b 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
y 1021 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
b 2119 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
y 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
b 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

ESULT 13
S-10-149-736-44
Sequence 44, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06958
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/311126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1

SEQ ID NO 44
LENGTH: 11443
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
S-10-149-736-44

Query Match . 52.5%; Score 1103.4; DB 15; Length 11443;
Best Local Similarity 99.9%; Pred. No. 6.6e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACACGGAGCCCAT 60
b 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACACGGAGCCCAT 1158
y 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
b 1159 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
y 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278

QY 181 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGTTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAAAGAACAAAGGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAAAGAACAAAGGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACAATAAGGTGCTTCA 600
Db 1639 GCCTCTTGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACAATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTTCAGA 840
Db 1879 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTTCAGA 1938
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
Db 2119 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 14
US-10-149-736-47
; Sequence 47, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47
LENGTH: 12057
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
IS-10-149-736-47

Query Match 52.5%; Score 1103.4; DB 15; Length 12057;
Best Local Similarity 99.9%; Pred. No. 6.8e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

iy	1	GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGTGACCCCTACACGGAGCCCAT	60
ib	1099	GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGTGACCCCTACACGGAGCCCAT	1158
iy	61	TCCTTCACAGCATTTGGAGCTCCTGAAGACAACTATTTGGCAGTTCATTGATGGAGAG	120
ib	1159	TCCTTCACAGCATTTGGAGCTCCTGAAGACAACTATTTGGCAGTTCATTGATGGAGAG	1218
iy	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTATCGTGGCTTCTTC	180
ib	1219	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTATCGTGGCTTCTTC	1278
iy	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGAAAGTGGTGAAGA	240
ib	1279	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGAAAGTGGTGAAGA	1338
iy	241	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	300
ib	1339	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	1398
iy	301	TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA	360
ib	1399	TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA	1458
iy	361	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC	420
ib	1459	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC	1518
iy	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	480
ib	1519	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	1578
iy	481	GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGAACAGGAAATGGAGGAAGA	540
ib	1579	GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGAACAGGAAATGGAGGAAGA	1638
iy	541	GCCTCTTGACCTGATCTTTGAGACCTTAAACCGCCCAAGTACAACACATAAGGTGCTTCA	600
ib	1639	GCCTCTTGACCTGATCTTTGAGACCTTAAACCGCCCAAGTACAACACATAAGGTGCTTCA	1698
iy	601	AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCACTCACATGGTGGTGTAGT	660
ib	1699	AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCACTCACATGGTGGTGTAGT	1758
iy	661	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATGGG	720
ib	1759	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATGGG	1818
iy	721	AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGTTCITTTACAAGACAT	780
ib	1819	AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGTTCITTTACAAGACAT	1878

QY	781	CCTTCTCAAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTCTAGTGCATGGCTTTT	840
Db	1879	CCTTCTCAAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTCTAGTGCATGGCTTTT	1938
QY	841	AAAAGAAAGATGCAGTGAACAAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	900
Db	1939	AAAAGAAAGATGCAGTGAACAAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	1998
QY	901	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT	960
Db	1999	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT	2058
QY	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACACTGAAGAAATAAGTCAGTGAC	1020
Db	2059	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACACTGAAGAAATAAGTCAGTGAC	2118
QY	1021	CCAGAAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	1080
Db	2119	CCAGAAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	2178
QY	1081	ACTTGAAGAGAGTACAGCACAGACT	1105
Db	2179	ACTTGAAGAGAGTACAGCACAGACT	2203

RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match 52.5%; Score 1103.4; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 7.5e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	60
Db	1099	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	1158
QY	61	TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG	120
Db	1159	TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG	1218
QY	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	180
Db	1219	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	1278
QY	181	TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	240
Db	1279	TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	1338
QY	241	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	300
Db	1339	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	1398
QY	301	TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTAATCAGAAGATGAAGA	360

b 1399 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATCAAGA 1458
y 361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 420
b 1459 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 1518
y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
b 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
y 481 GAAAGAGTTGAATGCACTGGCTAAACAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 540
b 1579 GAAAGAGTTGAATGCACTGGCTAAACAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 1638
y 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAAACAATAAGTGCTTCA 600
b 1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAAACAATAAGTGCTTCA 1698
y 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAAATTTCTCTCACTCACATGGTGGTGGTAGT 660
b 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAAATTTCTCTCACTCACATGGTGGTGGTAGT 1758
y 661 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 720
b 1759 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1818
y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
b 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
y 781 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 840
b 1879 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 1938
y 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
b 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
y 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
b 1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCACACTGAAGAATAAGTCAGTGAC 1020
b 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCACACTGAAGAATAAGTCAGTGAC 2118
y 1021 CCAGAAGACGGAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
b 2119 CCAGAAGACGGAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
y 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
b 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
M nucleic - nucleic search, using sw model
Run on: April 4, 2004, 11:55:32 ; Search time 542.141 Seconds
(without alignments)
16463.377 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Effect score: 2101
Sequence: 1 gagctatgctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		Description	
Result No.	Score	Match	Length	DB	ID		
1	2101	100.0	4182	6	AAD37230	Aad37230 Human dys	
2	2101	100.0	5149	6	AAD37255	Aad37255 Adeno-ass	
3	1579.8	75.2	3999	6	AAD37234	Aad37234 Human dys	
4	1579.8	75.2	4966	6	AAD37256	Aad37256 Adeno-ass	
5	1579.8	75.2	4990	6	AAD37262	Aad37262 Adeno-ass	
6	1443	68.7	3858	6	AAD37237	Aad37237 Human dys	
7	1443	68.7	4825	6	AAD37257	Aad37257 Adeno-ass	
8	1443	68.7	4848	6	AAD37263	Aad37263 Adeno-ass	
9	1443	68.7	5060	6	AAD37264	Aad37264 Adeno-ass	
10	1293	61.5	8689	6	ABK82000	Abk82000 DNA encod	
11	1131	53.8	5952	5	AAD06794	Aad06794 Human dys	
12	1114	53.0	4414	6	AAD37260	Aad37260 Adeno-ass	
13	1103.4	52.5	5417	6	ABK81997	Abk81997 DNA encod	
14	1103.4	52.5	11241	6	ABK82005	Abk82005 cDNA enco	
15	1103.4	52.5	11443	6	ABK82002	Abk82002 DNA encod	
16	1103.4	52.5	12923	1	AAN90338	Aan90338 Sequence	
17	1103.4	52.5	13957	6	ABK81959	Abk81959 cDNA enco	
18	1103.4	52.5	13957	6	ABT10904	Abt10904 Human bre	
19	1103.4	52.5	13957	6	ABN95786	Abn95786 Gene #228	
20	1103.4	52.5	13957	6	ABS69900	Abs69900 Human dys	
21	1103.4	52.5	13977	6	ABS70403	Abs70403 Human bon	
22	1102	52.5	3446	6	AAD37242	Aad37242 Human dys	
23	1100.2	52.4	11058	6	AAD37229	Aad37229 Human dys	

24	1097.8	52.3	1991	6	AAD37231	Aad37231 Human dys
25	999	47.5	2169	6	AAD37232	Aad37232 Human dys
26	999	47.5	3531	6	AAD37238	Aad37238 Human dys
27	999	47.5	4498	6	AAD37258	Aad37258 Adeno-ass
28	997	47.5	5339	6	ABK81998	Abk81998 DNA encod
29	996	47.4	5462	6	ABK81999	Abk81999 DNA encod
30	911.6	43.4	4402	3	AAZ48568	Aaz48568 A rod sho
31	866.6	41.2	13815	6	ABK81960	Abk81960 cDNA enco
32	866.6	41.2	13815	6	ABI99799	Abi99799 Mouse isc
33	866.6	41.2	19307	2	AAT27558	Aat27558 Shuttle v
34	858.6	40.9	13815	2	AAV18885	Aav18885 Mus muscu
35	855.4	40.7	3275	1	AAN97129	Aan97129 Partial s
36	787	37.5	3510	6	AAD37240	Aad37240 Human dys
37	787	37.5	4476	6	AAD37259	Aad37259 Adeno-ass
38	777	37.0	1667	6	AAD37235	Aad37235 Human dys
39	718.8	34.2	4075	3	AAZ48569	Aaz48569 A rod sho
40	714.6	34.0	4402	3	AAZ48567	Aaz48567 A rod sho
41	652	31.0	1821	6	AAD37241	Aad37241 Human dys
42	450	21.4	1340	6	AAD37239	Aad37239 Human dys
43	409.2	19.5	3747	3	AAZ48566	Aaz48566 A rod sho
44	387	18.4	387	6	ABK81990	Abk81990 Human dys
45	361.4	17.2	9195	7	ACD19399	Acd19399 cDNA enco

ALIGNMENTS

RESULT 1
AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 43-44; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
Query Match 100.0%; Score 2101; DB 6; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 959
61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
960 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1019
121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
1020 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1079
181 TGTCTAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1080 TGTCTAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAGTTTCATACTCATAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
1140 CCAGTTTCATACTCATAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1199
301 TAATATTCTACAATTGGGAAGTAAGCTGATTCCTTAATTCGAAGTGGGATGCCTCAGGGTAGC 360
1200 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTATCAGAAGATGAAGA 1259
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGTGGGATGCCTCAGGGTAGC 420
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGTGGGATGCCTCAGGGTAGC 1319
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCAGAAATCAGAAACT 480
1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCAGAAATCAGAAACT 1379
481 GAAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 540
1380 GAAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 1439
541 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
1440 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
1500 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGT 1559
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACCACTTAAGGTATTGGG 720
1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACCACTTAAGGTATTGGG 1619
721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 780
1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1679
781 CCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 840
1680 CCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1739
841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
1740 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1799
901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCCGATCTAGAAAAGAAAAAGCAATCCAT 960

Db 1800 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1859
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAAGTAC 1020
Db 1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAAGTAC 1919
Qy 1021 CCAGAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 1920 CCAGAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAAA 1979
Qy 1081 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGA 1140
Db 1980 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGA 2039
Qy 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
Db 2040 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTTACAGGATGCTAC 2099
Qy 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA 1260
Db 2100 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA 2159
Qy 1261 AGACCTCAAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320
Db 2160 AGACCTCAAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 2219
Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCCCTGTTACAAAAGACGTTT 1380
Db 2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCCCTGTTACAAAAGACGTTT 2279
Qy 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2280 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 2339
Qy 1441 TTTGGAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2340 TTTGGAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2399
Qy 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 1560
Db 2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 2459
Qy 1561 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGA 1620
Db 2460 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGA 2519
Qy 1621 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 2520 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAGCCTTTTGA 2579
Qy 1681 AGGACTAGAGAAAACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 1740
Db 2580 AGGACTAGAGAAAACTCTACAGGAGCCCAGAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 2639
Qy 1741 TGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAA 1800
Db 2640 TGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAA 2699
Qy 1801 CCGTCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 2700 CCGTCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 2759
Qy 1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2760 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 2819
Qy 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGAAAGTCAA 1980
Db 2820 CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGAAAGTCAA 2879
Qy 1981 GGCACCTCGAGGAGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040
Db 2880 GGCACCTCGAGGAGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2939

Y 2041 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 2100
b 2940 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 2999
Y 2101 C 2101
b 3000 C 3000

ESULT 2
AD37255
D AAD37255 standard; DNA; 5149 BP.
X AAD37255;
X
T 21-AUG-2002 (first entry)
X
E Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
X
W Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X

S Homo sapiens.
S Unidentified.
S Chimeric.
X WO200183695-A2.
X
D 08-NOV-2001.
X
F 27-APR-2001; 2001WO-US013677.
X
R 28-APR-2000; 2000US-0200777P.
X
X (XIAO/) XIAO X.
X
I Xiao X;
X
X WPI; 2002-049342/06.
X
T New dystrophin minigene for treating Duchenne or Becker muscular
T dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.

X Example 1; Page 57-59; 71pp; English.
X
S The present invention relates to an isolated nucleotide sequence encoding
S a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is
C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C dystrophy (BMD) in a mammalian subject. The present sequence is AAV
C vector plasmid construct containing human dystrophin minigenes, a muscle
C creatine kinase (MCK) promoter and a small polyA signal sequence

X Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;
Q
Query Match 100.0%; Score 2101; DB 6; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCATT 60
b 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCATT 1716
Y 61 TCCTTCACAGCATTGGGAAGCTCCTGAACACAAGTCATTTGSCAGTTTCATTGATGGAGAG 120

Db 1717 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGTTGG 300
Db 1897 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGTTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGACGAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGACGAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCATCATGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCATCATGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
Db 2377 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 2436
QY 781 CCTTCTCAAATGGCAAGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTTCAGA 840
Db 2437 CCTTCTCAAATGGCAAGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTTCAGA 2496
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 900
Db 2497 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 2556
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2616
QY 961 GGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2617 GGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2736
QY 1081 ACTTGAAAAGAGTACAGCAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
Db 2737 ACTTGAAAAGAGTACAGCAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2796
QY 1141 AAAGTTTCTTGGCTTACAGAAGCTGAACAACTGCCAATGTCTCTACAGGATGCTAC 1200

b 2797 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACACTGCCAATGTCCTACAGGATGCTAC 2856
y CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 1260
b 2857 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 2916
y AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAG 1320
b 2917 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAG 2976
y CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 1380
b 2977 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 3036
y GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCTCA 1440
b 3037 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCTCA 3096
y TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 1500
b 3097 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 3156
y GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
b 3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3216
y AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATTGAAAACTAAAGA 1620
b 3217 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATTGAAAACTAAAGA 3276
y ACCTGTAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAGCCTTTTGA 1680
b 3277 ACCTGTAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAGCCTTTTGA 3336
y AGGACTAGAGAAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAGCCAGAA 1740
b 3337 AGGACTAGAGAAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAGCCAGAA 3396
y TGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
b 3397 TGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3456
y CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860
b 3457 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 3516
y TCAAGAGCCACGGATGAGCTGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 1920
b 3517 TCAAGAGCCACGGATGAGCTGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 3576
y CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
b 3577 CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3636
y GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
b 3637 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3696
y TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGA 2100
b 3697 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGA 3756
y 2101 C 2101
b 3757 C 3757

RESULT 3
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
XX

DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3990.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW ageno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US013677.
XX 28-APR-2000; 2000US-0200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
PI
XX WPI; 2002-049342/06.
DR New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 46-47; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;
SQ
Query Match 75.2%; Score 1579.8; DB 6; Length 3999;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
Qy 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 60
Db 900 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 959
Qy 61 TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAGTCATTTGCGAGTTTCATTGATGGAGAG 120
Db 960 TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAGTCATTTGCGAGTTTCATTGATGGAGAG 1019
Qy 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1079
Qy 181 TGCTGAGGACACATTCGAAGCACAAGAGAGATTTCTATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTCGAAGCACAAGAGAGATTTCTATGATGTGGAAGTGGTGAAGA 1139
Qy 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACGCCCATCAGGGCCGGGTGG 300
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACGCCCATCAGGGCCGGGTGG 1199
Qy 301 TAAATATTCTACAATTTGGGAAGTAAAGCTGATTTGGAAACAGGAAATTTATCAGAAGATGAAGA 360

b 1200 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1259
y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 420
b 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 1319
y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT 480
b 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT 1379
y 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
c 1380 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1439
y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAACATAGGTCCTTCA 600
b 1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAACATAGGTCCTTCA 1499
y 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCCTCACTCACATGGTGGGTAGT 660
b 1500 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCCTCACTCACATGGTGGGTAGT 1559
y 661 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
b 1560 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1619
y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
b 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACCA 1679
y 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTTCA 840
b 1680 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 1727
y 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 900
b 1728 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 1777
y 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
b 1778 ----- 1777
y 961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACITGAAGAATAAGTCAGTGAC 1020
b 1778 -----CCAAACTAGAATGCCATCTTCTCCTGATGTTGGAG----- 1812
y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
b 1813 ----- 1812
y 1081 ACTTGAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
b 1813 -----GTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1856
y 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACTGCCAATGTCCTACAGGATGCTAC 1200
b 1857 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACTGCCAATGTCCTACAGGATGCTAC 1916
y 1201 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
b 1917 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1976
y 1261 AGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAG 1320
b 1977 AGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAG 2036
y 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 1380
b 2037 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 2096
y 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA 1440
b 2097 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA 2156

QY 1441 TTTGGAAGCCAGTTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
Db 2157 TTTGGAAGCCAGTTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 2216
QY 1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCC 1560
Db 2217 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCC 2276
QY 1561 AGCAGTTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTTGAAAACTAAAGA 1620
Db 2277 AGCAGTTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTTGAAAACTAAAGA 2336
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATATTTCTGACAGAGCAGCCTTTGGA 1680
Db 2337 ACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATATTTCTGACAGAGCAGCCTTTGGA 2396
QY 1681 AGGACTAGAGAACTCTTACCAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1740
Db 2397 AGGACTAGAGAACTCTTACCAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 2456
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
Db 2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 2516
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 1860
Db 2517 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 2576
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2577 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 2636
QY 1921 CTGCGAGCCCGTGGCGGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAA 1980
Db 2637 CTGCGAGCCCGTGGCGGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAA 2696
QY 1981 GGCACCTTCAGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCACCTTGC 2040
Db 2697 GGCACCTTCAGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCACCTTGC 2756
QY 2041 TCGCAGCTTACCACCTTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTTGAAGA 2100
Db 2757 TCGCAGCTTACCACCTTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTTGAAGA 2816
QY 2101 C 2101
Db 2817 C 2817

RESULT 4
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX

AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX

DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

OS Homo sapiens.
OS Unidentified.
OS Chimeric.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.

28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
Example 1; Page 59-60; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence
Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;
Query Match 75.2%; Score 1579.8; DB 6; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1716
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1776
121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 180
1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 1836
181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1837 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTTGG 300
1897 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTTGG 1956
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 360
1957 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 2016
361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
2017 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076
421 TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
2077 TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAATGAGGGAAGA 540
2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAATGAGGGAAGA 2196
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACACATAGGTGCTTCA 600

2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAAGTGTCTTCA 2256
601 AGAAGATCTAGAAACAAGAACCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
2257 AGAAGATCTAGAAACAAGAACCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2316
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 720
2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 2376
721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAAGACAT 780
2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAAGACCA 2436
781 CCTTCTCAAAATGGCAACGTCCTTACTGAAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 840
2437 GCCTGAC-----CTAGTCTCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2484
841 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAATGTT 900
2485 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2534
901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
2535 ----- 2534
961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
2535 -----CCAAACTAGAAAATGCCATCTTCTCTGTGATGTTGGAG----- 2569
1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
2570 ----- 2569
1081 ACTTGAAGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 1140
2570 -----GTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 2613
1141 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACCTGCCAATGTCTCTACAGGATGCTAC 1200
2614 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAAACAACCTGCCAATGTCTCTACAGGATGCTAC 2673
1201 CCGTAAGGAAAGGCTCCTAGAACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
2674 CCGTAAGGAAAGGCTCCTAGAACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2733
1261 AGACCTCCAAAGTGAAATTTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAACAG 1320
2734 AGACCTCCAAAGTGAAATTTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAACAG 2793
1321 CCAAAAAATCCTGAGATCCCTTGAAGGTTCCGATGATGCAAGTCTGTGTTACAAAAGACGTTT 1380
2794 CCAAAAAATCCTGAGATCCCTTGAAGGTTCCGATGATGCAAGTCTGTGTTACAAAAGACGTTT 2853
1381 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTTAGGTCCCA 1440
2854 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTTAGGTCCCA 2913
1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGT 1500
2914 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGT 2973
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGGCGACTTTC 1560
2974 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGGCGACTTTC 3033
1561 AGCAGTTTCAGAACAGAACGATGATGATAGGGCCTTCAAGAGGGAATTTGAAAACCTTAAAGA 1620
3034 AGCAGTTTCAGAACAGAACGATGATGATAGGGCCTTCAAGAGGGAATTTGAAAACCTTAAAGA 3093
1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
3094 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 3153

1681 AGGACTAGAGAAACTCTACGAGAGCCAGAGAGCTGCTCCTCTGAGGAGAGCCAGAA 1740
|||||
3154 AGGACTAGAGAAACTCTACGAGAGCCAGAGAGCTGCTCCTCTGAGGAGAGCCAGAA 3213
|||||
1741 TGTCACTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA 1800
|||||
3214 TGTCACTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA 3273
|||||
1801 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 1860
|||||
3274 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACT 3333
|||||
1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 1920
|||||
3334 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 3393
|||||
1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
|||||
3394 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3453
|||||
1981 GGCACTTCGAGAGAAATTGGCCCTCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
|||||
3454 GGCACTTCGAGAGAAATTGGCCCTCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3513
|||||
2041 TCGCCAGCTTACACATTTGGGCAATTCAGCTCTCAACGTAATACCTCAGCACTCTGGAAGA 2100
|||||
3514 TCGCCAGCTTACACATTTGGGCAATTCAGCTCTCAACGTAATACCTCAGCACTCTGGAAGA 3573
|||||
2101 C 2101
3574 C 3574

RESULT 5
AD37262
AAD37262 standard; DNA; 4990 BP.
AAD37262;
21-AUG-2002 (first entry)
Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
Homo sapiens.
Cytomegalovirus.
Unidentified.
Chimeric.
WO200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.
Example 1; Page 67-68; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomagalovirus (CMV) promoter and a small polyA signal sequence
XX Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;
SQ
Query Match 75.2%; Score 1579.8; DB 6; Length 4990;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
|||||
Db 1681 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1740
|||||
QY 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
|||||
Db 1741 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1800
|||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
|||||
Db 1801 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1860
|||||
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
Db 1861 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1920
|||||
QY 241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
Db 1921 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1980
|||||
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 360
|||||
Db 1981 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 2040
|||||
QY 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 2041 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2100
|||||
QY 421 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
Db 2101 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAATCAGAAACT 2160
|||||
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 540
|||||
Db 2161 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 2220
|||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAGGTGCTTCA 600
|||||
Db 2221 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAGGTGCTTCA 2280
|||||
QY 601 AGAAGATCTAGAACAAGAAACAAGTCAAGGTCAATTTCTCACTCACAATGGTGGTAGT 660
|||||
Db 2281 AGAAGATCTAGAACAAGAAACAAGTCAAGGTCAATTTCTCACTCACAATGGTGGTAGT 2340
|||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
|||||
Db 2341 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2400
|||||
QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
|||||
Db 2401 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACCA 2460
|||||
QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTTCA 840
|||||

2461 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2508
841 AAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAAATGTT 900
2509 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2558
901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
2559 ----- 2558
961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1020
2559 -----CCAAACTAGAAATGCCATCTTCTTGATGTTGGAG----- 2593
1021 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
2594 ----- 2593
1081 ACTTGAAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
2594 -----GTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2637
1141 AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 1200
2638 AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 2697
1201 CCGTAAGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
2698 CCGTAAGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2757
1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTTGGATGAAAACAG 1320
2758 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTTGGATGAAAACAG 2817
1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAAGACGTTT 1380
2818 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAAGACGTTT 2877
1381 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
2878 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTAGGTCCCA 2937
1441 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGACCTTTCTCTGAGGAACTTCTGGT 1500
2938 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGACCTTTCTCTGAGGAACTTCTGGT 2997
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCC 1560
2998 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCC 3057
1561 AGCAGTTCAGAAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGGAAATTGAAAACTAAAGA 1620
3058 AGCAGTTCAGAAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGGAAATTGAAAACTAAAGA 3117
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
3118 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3177
1681 AGGACTAGAGAAAACTCTACCAGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGAA 1740
3178 AGGACTAGAGAAAACTCTACCAGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGAA 3237
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
3238 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3297
1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGACTCCAGGAAC 1860
3298 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGACTCCAGGAAC 3357
1861 TCAAGAGGCGACGGATGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGATCAAGGGATC 1920

Db 3358 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 3417
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 3418 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3477
QY 1981 GGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGC 2040
Db 3478 GGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGC 3537
QY 2041 TCGCCAGCTTACCACCTTTGGGATTGAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 2100
Db 3538 TCGCCAGCTTACCACCTTTGGGATTGAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 3597
QY 2101 C 2101
Db 3598 C 3598
RESULT 6
AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;
Query Match 68.7%; Score 1443; DB 6; Length 3858;

Best Local Similarity 84.6%; Pred. No. 0;					Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;				
Y	1	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAATT	60						
b	900	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAATT	959						
Y	61	TCCTTCACAGCAATTGGAAAGCTCCTGAAGACAAAGTCAATTTGGGAGTTTCATTGTATGGAGAG	120						
b	960	TCCTTCACAGCAATTGGAAAGCTCCTGAAGACAAAGTCAATTTGGGAGTTTCATTGTATGGAGAG	1019						
Y	121	TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC	180						
b	1020	TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC	1079						
Y	181	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGTGAAAGTGGTGAAGA	240						
b	1080	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGTGAAAGTGGTGAAGA	1139						
Y	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	300						
b	1140	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	1199						
Y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTTGAAACAGGAAATATATCAGAAGATGAAGA	360						
b	1200	TAATATTCTACAATTGGGAAGTAAGCTGATTTGAAACAGGAAATATATCAGAAGATGAAGA	1259						
Y	361	AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	420						
b	1260	AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	1319						
Y	421	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT	480						
b	1320	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT	1379						
Y	481	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGGAAATGGAGGAAGA	540						
b	1380	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGGAAATGGAGGAAGA	1439						
Y	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTGCTTCA	600						
b	1440	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTGCTTCA	1499						
Y	601	AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT	660						
b	1500	AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT	1559						
Y	661	TGATGAATCTAGTGGAGATCAGCGCACTGCTGCTTTGGAGAAACAACCTTAAGGTATTGGG	720						
b	1560	TGATGAATCTAGTGGAGATCAGCGCACTGCTGCTTTGGAGAAACAACCTTAAGGTATTGGG	1619						
Y	721	AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT	780						
b	1620	AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGAC--	1677						
Y	781	CCTTCTCAAATGGCAACGTCTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA	840						
b	1678	-----	1677						
Y	841	AAAAGAAGATGCAAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT	900						
b	1678	-----	1677						
Y	901	ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	960						
b	1678	-----	1677						
Y	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAAGTGC	1020						
b	1678	-----	1677						
Y	1021	CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA	1080						

Db	1678	-----	1677
Qy	1081	ACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA	1140
Db	1678	-----	1675
Qy	1141	AAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC	1200
Db	1716	AAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC	1775
Qy	1201	CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGTGATGATAAACAATGGCA	1260
Db	1776	CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGTGATGATAAACAATGGCA	1835
Qy	1261	AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTTGGATGAAAACAG	1320
Db	1836	AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTTGGATGAAAACAG	1895
Qy	1321	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT	1380
Db	1896	CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT	1955
Qy	1381	GGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAATAAAGTCTCTCAACATTAGGTCCCA	1440
Db	1956	GGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAATAAAGTCTCTCAACATTAGGTCCCA	2015
Qy	1441	TTTGGAAAGCCAGTTCTGACCAAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT	1500
Db	2016	TTTGGAAAGCCAGTTCTGACCAAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT	2075
Qy	1501	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCC	1560
Db	2076	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCC	2135
Qy	1561	AGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA	1620
Db	2136	AGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA	2195
Qy	1621	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA	1680
Db	2196	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA	2255
Qy	1681	AGGACTAGAGAAACTCTACAGGAGGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA	1740
Db	2256	AGGACTAGAGAAACTCTACAGGAGGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA	2315
Qy	1741	TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA	1800
Db	2316	TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA	2375
Qy	1801	CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC	1860
Db	2376	CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC	2435
Qy	1861	TCAAAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC	1920
Db	2436	TCAAAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC	2495
Qy	1921	CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	1980
Db	2496	CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA	2555
Qy	1981	GGCAGTTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC	2040
Db	2556	GGCAGTTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC	2615
Qy	2041	TCGCCAGCTTACCAGTTTGGGCATTCAGTCTCTCACCGTATAACCTCAGCACTCTGGAAGA	2100
Db	2616	TCGCCAGCTTACCAGTTTGGGCATTCAGTCTCTCACCGTATAACCTCAGCACTCTGGAAGA	2675
Qy	2101	C 2101	
Db	2676	C 2676	

RESULT 7

AD37257
ID AAD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.

XX Homo sapiens.
XX Unidentified.
XX Chimeric.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.

XX 28-APR-2000; 2000US-020077P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.

XX Example 1; Page 61-62; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) promoter and a small polyA signal sequence

XX Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

XX Query Match 68.7%; Score 1443; DB 6; Length 4825;

XX Best Local Similarity 84.6%; Pred. No. 0;

XX Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

XX 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 60
XX
XX 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 1716

XX 61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
XX
XX 1717 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776

XX 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
XX
XX 1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

XX 181 TGCTGAGGACACATTGCAAGACAAGGAGAGATTCTTCTAATGATGTGGAAGTGGTGAAGA 240

Db 1837 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATFACTCATGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1897 CCAGTTTCATFACTCATGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGGTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGCTAGC 420
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGCTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAAACT 480
Db 2077 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTCTTTTACAGACAT 780
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTCTTTTACAGAC-- 2434
QY 781 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 840
Db 2435 ----- 2434
QY 841 AAAAGAAGATGCAGTGAAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 2435 ----- 2434
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 2435 ----- 2434
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2435 ----- 2434
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAAATTTAGTCCAAAA 1080
Db 2435 ----- 2434
QY 1081 ACTTGAAGAAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
Db 2435 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2472
QY 1141 AAAAGTTTCTTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTACAGGATGCTAC 1200
Db 2473 AAAAGTTTCTTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTACAGGATGCTAC 2532
QY 1201 CCGTAAGGAAAAGGCTCCTAGAGAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 2533 CCGTAAGGAAAAGGCTCCTAGAGAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2592
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320

b 2593 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAG 2652
Y 1321 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTGTACAAAAGCGTTT 1380
b 2653 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTGTACAAAAGCGTTT 2712
Y 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
b 2713 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2772
Y 1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
b 2773 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2832
Y 1501 GTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGCGGACTTCC 1560
b 2833 GTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGCGGACTTCC 2892
Y 1561 AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGA 1620
b 2893 AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGA 2952
Y 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCCTTGGGA 1680
b 2953 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCCTTGGGA 3012
Y 1681 AGGACTAGAGAAACTCTACCAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
b 3013 AGGACTAGAGAAACTCTACCAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3072
Y 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
b 3073 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3132
Y 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGCTCCAGGAACT 1860
b 3133 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGCTCCAGGAACT 3192
Y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
b 3193 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3252
Y 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
b 3253 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3312
Y 1981 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040
b 3313 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3372
Y 2041 TCGCCAGCTTACCACCTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACCTCTGGAAGA 2100
b 3373 TCGCCAGCTTACCACCTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACCTCTGGAAGA 3432
Y 2101 C 2101
b 3433 C 3433

ESULT 8

AD37263
D AAD37263 standard; DNA; 4848 BP.

X AAD37263;

X 21-AUG-2002 (first entry)

T Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.

OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 68-70; 71pp; English.
XX

CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match 68.7%; Score 1443; DB 6; Length 4848;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1680 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1739
QY 61 TCCTTACAGCATTTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1740 TCCTTACAGCATTTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1799
QY 121 TGAAGTAAACCTGGACCCGTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 180
Db 1800 TGAAGTAAACCTGGACCCGTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 1859
QY 181 TGCTGAGGACACATTGCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1860 TGCTGAGGACACATTGCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1919
QY 241 CCAGTTTCATACACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
Db 1920 CCAGTTTCATACACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 1979
QY 301 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1980 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2039
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2040 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2099

2Y 421 TAGCATGGAAAAACAAAGCAATTATACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
2b 2100 TAGCATGGAAAAACAAAGCAATTATACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2159
2Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAACAAAGGAAATGGAGGAAGA 540
2b 2160 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAACAAAGGAAATGGAGGAAGA 2219
2Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAAGGTGCTTCA 600
2b 2220 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAAGGTGCTTCA 2279
2Y 601 AGAAGATCTAGAAACAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
2b 2280 AGAAGATCTAGAAACAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2339
2Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
2b 2340 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2399
2Y 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
2b 2400 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGAC-- 2457
2Y 781 CCTTCTCAATGGCAACGCTTACTGAAGAACAGTCAGTGCTTTTTAGTGCAATGGCTTTCAGA 840
2b 2458 ----- 2457
2Y 841 AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
2b 2458 ----- 2457
2Y 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
2b 2458 ----- 2457
2Y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
2b 2458 ----- 2457
2Y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAATTAGTCCAAAA 1080
2b 2458 ----- 2457
2Y 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
2b 2458 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2495
2Y 1141 AAAGTTTCTTCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTACAGGATGCTAC 1200
2b 2496 AAAGTTTCTTCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTACAGGATGCTAC 2555
2Y 1201 CCGTAAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAACAATGGCA 1260
2b 2556 CCGTAAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAACAATGGCA 2615
2Y 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320
2b 2616 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 2675
2Y 1321 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGAGTCCCTGTTTACAAAAGACGTTT 1380
2b 2676 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGAGTCCCTGTTTACAAAAGACGTTT 2735
2Y 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 1440
2b 2736 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 2795
2Y 1441 TTTGGAAGCCACTTCTGACCAGTGGAGCGTCTGCACCTTCTCTGCAGGAAGTCTCTGGT 1500
2b 2796 TTTGGAAGCCAGTCTTGACCAGTGGAGCGTCTGCACCTTCTCTGCAGGAAGTCTCTGGT 2855
2Y 1501 GTGGCTACAGTGAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTC 1560

Db 2856 GTGGCTACAGTGAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGACTTTC 2915
QY 1551 AGCAGTTTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 2916 AGCAGTTTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2975
QY 1621 ACCTGTAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
Db 2976 ACCTGTAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 3035
QY 1681 AGGACTAGAGAAACTCTACCAGGAGCCCGCAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAA 1740
Db 3036 AGGACTAGAGAAACTCTACCAGGAGCCCGCAGAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAA 3095
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
Db 3096 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3155
QY 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGAGACCCCTTGAAGACTCCAGGAAC 1860
Db 3156 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGAGACCCCTTGAAGACTCCAGGAAC 3215
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
Db 3216 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 3275
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 3276 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3335
QY 1981 GGCACCTTCGAGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 2040
Db 3336 GGCACCTTCGAGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGC 3395
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGA 2100
Db 3396 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGA 3455
QY 2101 C 2101
Db 3456 C 3456

RESULT 9
AAD37264

ID AAD37264 standard; DNA; 5060 BP.

AC AAD37264;

DT 21-AUG-2002 (first entry)

DE Adeno-associated virus (AAV) vector plasmid, AAV-B-CMV-3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.

OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.

PN WO200183695-A2.

XX 08-NOV-2001.

PF 27-APR-2001; 2001WO-US013677.

PR 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

XX

I Xiao X;
X WPI; 2002-049342/06.
R New dystrophin minigene for treating Duchenne or Becker muscular
X dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.
T
X
S Example 1; Page 70-71; 71pp; English.
X
C The present invention relates to an isolated nucleotide sequence encoding
C a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is
C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C dystrophy (BMD) in a mammalian subject. The present sequence is AAV
C vector plasmid construct containing human dystrophin minigenes, a muscle
C creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
C small polyA signal sequence
X
Q Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Query Match 68.7%; Score 1443; DB 6; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 60
b |||||
Y 1892 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 1951
b |||||
Y 61 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAGTCATTTGSCAGTTTCATTGTATGAGAGAG 120
b |||||
Y 1952 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAGTCATTTGSCAGTTTCATTGTATGAGAGAG 2011
b |||||
Y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
b |||||
Y 2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 2071
b |||||
Y 181 TGCTGAGGACACATTCGAAGCACAAGGACAGATTTCTAATGATGTGGAACTGGTGAAGA 240
b |||||
Y 2072 TGCTGAGGACACATTCGAAGCACAAGGACAGATTTCTAATGATGTGGAACTGGTGAAGA 2131
b |||||
Y 241 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTGACAGGCCCATCAGGGCCGGGTTGG 300
b |||||
Y 2132 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTGACAGGCCCATCAGGGCCGGGTTGG 2191
b |||||
Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
b |||||
Y 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2251
b |||||
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
b |||||
Y 2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2311
b |||||
Y 421 TAGCATGGAAAAACAAGCAATTTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 480
b |||||
Y 2312 TAGCATGGAAAAACAAGCAATTTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 2371
b |||||
Y 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA 540
b |||||
Y 2372 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA 2431
b |||||
Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAAATAGGTGCTTCA 600
b |||||
Y 2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAAATAGGTGCTTCA 2491
b |||||
Y 601 AGAAGATCTAGAACAAAGAAACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
b |||||

Db 2492 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAANTTCTCTCACTCACATGGTGGTGTAGT 2551
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2611
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 780
Db 2612 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGAC-- 2669
QY 781 CCITCTCAAATGGCAACGCTTACTGAAGAACAGTCGCTTTTGTAGTCATGGCTTTTCAGA 840
Db 2670 ----- 2669
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACACTGCTGGCTTTAAAGATCAAAAATGAAATGTT 900
Db 2670 ----- 2669
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 2670 ----- 2669
QY 961 GGGCAAACTGTATTCTCACTCAAAACAAGATCTTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
Db 2670 ----- 2669
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
Db 2670 ----- 2669
QY 1081 ACTTGAAGAAGATACAGCACAGACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 1140
Db 2670 -----ACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 2707
QY 1141 AAAGTTTCTTGCTGCTTACAGAAGCTGAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
Db 2708 AAAGTTTCTTGCTGCTTACAGAAGCTGAACAACTGCCAATGTCTCTACAGGATGCTAC 2767
QY 1201 CCGTAAGGAAAGGCTCCTAGAAGA CTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 2768 CCGTAAGGAAAGGCTCCTAGAAGA CTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 2827
QY 1261 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
Db 2828 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2887
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT 1380
Db 2888 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT 2947
QY 1381 GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2948 GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAAGTCTCTCAACATTAGGTCCCA 3007
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
Db 3008 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 3067
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 3068 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 3127
QY 1561 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAAAACTAAAAGA 1620
Db 3128 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAAAACTAAAAGA 3187
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 3188 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3247
QY 1681 AGGACTAGAGAAAACCTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
Db 3248 AGGACTAGAGAAAACCTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3307

2Y 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAATTGAA 1800
|||
Db 3308 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAATTGAA 3367
|||
2Y 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
|||
Db 3368 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3427
|||
2Y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
|||
Db 3428 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 3487
|||
2Y 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
|||
Db 3488 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3547
|||
2Y 1981 GGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCACGTCGAATGACCTTGC 2040
|||
Db 3548 GGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCACGTCGAATGACCTTGC 3607
|||
2Y 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
|||
Db 3608 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3667
|||
2Y 2101 C 2101
|
Db 3668 C 3668

RESULT 10
ID ABK82000 standard; DNA; 8689 BP.
CX
AC ABK82000;
CX
JT 13-AUG-2002 (first entry)
CX
DE DNA encoding mini-dystrophin protein deltaH2-R19.
CX
CW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
CX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
CX
DS Homo sapiens.
DS Synthetic.
CX
PN WO200229056-A2.
CX
CX 11-APR-2002.
CX
CX 04-OCT-2001; 2001WO-US031126.
CX
CX 06-OCT-2000; 2000US-0238848P.
CX
CX (UNMI) UNIV MICHIGAN.
CX
CX Chamberlain JS, Harper SQ;
CX
CX WPI; 2002-435334/46.
CX
CX A composition for preparing therapeutic drugs, has a mini-dystrophin
CX peptide comprising a specific number of spectrin-like repeat domains, or
CX a nucleic acid sequence encoding the mini-dystrophin peptide.
CX
CX Disclosure; Fig 15; 145pp; English.
CX
CX The invention describes a composition comprising a mini-dystrophin
CX peptide comprising a spectrin-like repeat domain, where the domain
CX comprises n spectrin-like repeats, and contains no more than n spectrin-
CX like repeats, where n is an even number between 4-24, or a nucleic acid
CX encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CX polynucleotide encoding it is useful as a medicament, for preparing a
CX drug for therapeutic application and in the preparation of a composition

CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;
Query Match 61.5%; Score 1293; DB 6; Length 8689;
Best Local Similarity 72.5%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;
QY 1 GAGCTATGCTACACACAGGCTGTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
|||
Db 1099 GAGCTATGCTACACACAGGCTGTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158
|||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
|||
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
|||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
|||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
|||
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
|||
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 300
|||
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 1398
|||
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
|||
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
|||
QY 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 1459 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
|||
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
|||
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
|||
QY 481 GAAAGAGTTGAATGACTGGCTACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 540
|||
Db 1579 GAAAGAGTTGAATGACTGGCTACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 1638
|||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTCAATTTCTCACTCACATAGGTGCTTCA 600
|||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTCAATTTCTCACTCACATAGGTGCTTCA 1698
|||
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
|||
Db 1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 1758
|||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAAACAATTAAGGTATTGGG 720
|||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAAACAATTAAGGTATTGGG 1818
|||
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACGCGCTGGGTTCTTTTACAAGACAT 780
|||
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACGCGCTGGGTTCTTTTACAAGACAT 1878
|||
QY 781 CCTTCTCAAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTTCAGA 840
|||
Db 1879 CCTTCTCAAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTTCAGA 1938
|||
QY 841 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
|||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1998
|||
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGGATCTAGAAAAAGAAAAAGCAATCCAT 960
|||
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
|||

OX Homo sapiens.
 OX Key
 OX misc_recomb
 OX Location/Qualifiers
 OX 2847..2848
 OX /*tag= a
 OX /label= S4_junction_site
 OX /note= "Dystrophin gene is split at this site and
 OX attached to an intein sequence. The resulting fragment is
 OX used to produce plasmid pSD4"
 OX 2952..2953
 OX /*tag= b
 OX /label= S3_junction_site
 OX /note= "Dystrophin gene is split at this site and
 OX attached to an intein sequence. The resulting fragment is
 OX used to produce plasmid pSD3"
 OX 3198..3199
 OX /*tag= c
 OX /label= S2_junction_site
 OX /note= "Dystrophin gene is split at this site and
 OX attached to an intein sequence. The resulting fragment is
 OX used to produce plasmid pSD2"
 OX 3300..3301
 OX /*tag= d
 OX /label= S1_junction_site
 OX /note= "Dystrophin gene is split at this site and
 OX attached to an intein sequence. The resulting fragment is
 OX used to produce plasmid pSD1"
 OX
 OX WO200129243-A1.
 OX
 OX 26-APR-2001.
 OX
 OX 13-OCT-2000; 2000WO-CA001216.
 OX
 OX 15-OCT-1999; 99US-0159868P.
 OX
 OX (UYDA-) UNIV DALHOUSIE.
 OX (UYPI-) UNIV PITTSBURGH.
 OX
 OX Paul XL, Xiao X;
 OX WPI; 2001-367297/38.
 OX
 OX Use of spontaneous or automatic protein splicing to join two or more
 OX peptides at junction site involves expressing extein peptides having co-
 OX reacting portions of split intein attached to them, so that peptides
 OX splice.
 OX
 OX Example 1; Fig 2; 81pp; English.
 OX
 OX The invention relates to a method directed to the use of spontaneous or
 OX automatic protein trans-splicing to join two or more peptides at junction
 OX site. This method involves expressing extein peptides having co-reacting
 OX portions of split intein attached to them, so that peptides will splice
 OX automatically under suitable conditions. The invention also provides
 OX methods for circumventing virion packaging size limitations in
 OX recombinant virus particle, by splitting a coding region for a protein to
 OX be delivered into two or more extein genes, which are packaged in
 OX separate virus particles and are co-delivered in a target cell for the
 OX expression and for subsequent trans-splicing to form the complete
 OX protein. In particular, the method is used for trans-splicing human
 OX dystrophin and in gene therapies of recombinant adeno-associated virus
 OX (AAV) particles that encode trans-spliced dystrophin, for treating
 OX diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
 OX dystrophy (BMD). The present sequence is human dystrophin gene (Becker
 OX form)
 OX
 OX Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;
 OX
 OX Query Match 53.8%; Score 1131; DB 5; Length 5952;
 OX Best Local Similarity 68.6%; Pred. NO. 1.1e-306;
 OX Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACCGAGCCCAT 60
 Db |||||
 QY 891 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACCGAGCCCAT 950
 Db |||||
 QY 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
 Db |||||
 QY 951 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1010
 Db |||||
 QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 180
 Db |||||
 QY 1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 1070
 Db |||||
 QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
 Db |||||
 QY 1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1130
 Db |||||
 QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGTGG 300
 Db |||||
 QY 1131 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGTGG 1190
 Db |||||
 QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 360
 Db |||||
 QY 1191 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 1250
 Db |||||
 QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
 Db |||||
 QY 1251 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310
 Db |||||
 QY 421 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAA 480
 Db |||||
 QY 1311 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAA 1370
 Db |||||
 QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 540
 Db |||||
 QY 1371 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 1430
 Db |||||
 QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
 Db |||||
 QY 1431 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1490
 Db |||||
 QY 601 AGAAGATCTAGAAACAAGCAAGTCAAGGTCAATCTCTCACTCACATGGTGGTGGTAGT 660
 Db |||||
 QY 1491 AGAAGATCTAGAAACAAGCAAGTCAAGGTCAATCTCTCACTCACATGGTGGTGGTAGT 1550
 Db |||||
 QY 661 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 720
 Db |||||
 QY 1551 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 1610
 Db |||||
 QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAT 780
 Db |||||
 QY 1611 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAT 1670
 Db |||||
 QY 781 CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 840
 Db |||||
 QY 1671 CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 1730
 Db |||||
 QY 841 AAAAGAAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
 Db |||||
 QY 1731 AAAAGAAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1790
 Db |||||
 QY 901 ATCAAGTCTTCAAAAACTGGCGTTTTTAAAGCGGATCTAGAAAAAGAAAAACCAATCCAT 960
 Db |||||
 QY 1791 ATCAAGTCTTCAAAAACTGGCGTTTTTAAAGCGGATCTAGAAAAAGAAAAACCAATCCAT 1850
 Db |||||
 QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCTCAACACTGAAGAATAAGTCAGTGAC 1020
 Db |||||
 QY 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCTCAACACTGAAGAATAAGTCAGTGAC 1910
 Db |||||
 QY 1021 CCAGAAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
 Db |||||
 QY 1911 CCAGAAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1970
 Db |||||

1081 ACTTGAAGAGTACAGCACAG----- 1102
1971 ACTTGAAGAGTACAGCACAGGAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGT 2030
1103 ----- 1102
2031 GGAAGAGATTTGTCTAAAGGCGACGATTTGTACAAGGAAAAACCGCCACTCAGCCAGT 2090
1103 ----- 1102
2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAAGGCGGTAAACCGTTTACTTCAAGA 2150
1103 ----- 1102
2151 GCTAGGGGAAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210
1103 ----- 1102
2211 TCAGACTGTTACTCTGGTGACACAACTGTGTGTTACTAAGGAAACTGCCATCTCCAAACT 2270
1103 ----- 1102
2271 AGAAATGCCATCTTCTTGATGTTGGAGGTACCTGTCTGGCAGATTTCAACCGGGCTTG 2330
1103 ----- 1102
2331 GACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTATAAAATCACAGAGGTGAT 2390
1103 ----- 1102
2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGA 2450
1103 ----- 1102
2451 TTTGGAACAGAGCGCTCCCGAGTTGGAAGAACTCATTTACCGCTGCCCAAAATTTGAAAAA 2510
1103 ----- 1102
2511 CAAGACCAGCAATCAAGAGGCTAGAACAACTATTACGGATCGAATTGAAAGAAATTCAGAA 2570
1103 ----- 1102
2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGSCAACAGTTGAATGAAATGTT 2630
1103 ----- 1102
2631 AAAGGATTCACACAAATGGCTGGAAGCTAAGGAAGCTGAGCAGGTCTTAGCACAGGC 2690
1103 ----- 1102
2691 CAGAGCCAAGCTTGAGTCAATGGAAGAGGGTCCCTATACAGTAGATGCAATCCAAAAGAA 2750
1103 ----- 1102
2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGT 2810
1103 ----- 1102
2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGATTATTCTGCAGATGATACAGAAAAAGT 2870
1103 ----- 1102
2871 CCACATGATAACAGAGAATATCAATGCTCTTGGAGAAGCATTCATAAAAGGTTAGTGA 2930
1103 ----- 1102
2931 GCGAGAGGCTGCTTTGGAAGAACTCATAGATTACTGCAACAGTTCCCGCTGGACCTGGA 2990
1141 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
2991 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 3050
1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCCAAGGGAGTAAAGAGCTGTATGAAACAATGGCA 1260

3051 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 3110
1261 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAAAACAG 1320
3111 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAAAACAG 3170
1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTGTACAAAGACGTTT 1380
3171 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTGTACAAAGACGTTT 3230
1381 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
3231 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3290
1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGGCTTGACACCTTTCTCTGCAGGAACTTCTGGT 1500
3291 TTTGGAAGCCAGTTCTGACCAGTGGAAAGGCTTGACACCTTTCTCTGCAGGAACTTCTGGT 3350
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3410
1561 ACCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAACTAAAGA 1620
3411 ACCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAACTAAAGA 3470
1621 ACCTGTAATCATGAGTACTCTTGAGACITGTACGAATATTTCTGCAGAGCAGCCTTTTGA 1680
3471 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCTTTTGA 3530
1681 AGGACTAGAGAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
3531 AGGACTAGAGAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3590
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAA 1800
3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAA 3650
1801 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860
3651 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 3710
1861 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 1920
3711 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 3770
1921 CTGGCAGCCCGTGGCGGATCTCTCTCAATGACTCTCTCCAAGATCACTCGAGAAAGTCAA 1980
3771 CTGGCAGCCCGTGGCGGATCTCTCTCAATGACTCTCTCCAAGATCACTCGAGAAAGTCAA 3830
1981 GGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040
3831 GGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3890
2041 TCGCCAGCTTACCAGTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAGA 2100
3891 TCGCCAGCTTACCAGTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAGA 3950
2101 C 2101
3951 C 3951

RESULT 12
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX AAD37260;
XX AC
XX AC
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3447.

X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X
S Homo sapiens.
S Unidentified.
S Chimeric.
X WO200183695-A2.
N
X
D 08-NOV-2001.
X
F 27-APR-2001; 2001WO-US013677.
X
R 28-APR-2000; 2000US-0200777P.
X
A (XIAO/) XIAO X.
X
I Xiao X;
X
R WPI; 2002-049342/06.
X
T New dystrophin minigene for treating Duchenne or Becker muscular
T dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.
X
S Example 1; Page 65-66; 71pp; English.
X
X The present invention relates to an isolated nucleotide sequence encoding
X a dystrophin minigene. The minigene comprises N-terminal or modified N-
X terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
X domains and cysteine-rich domains of dystrophin or utrophin genes. The
X invention also relates to a recombinant adeno-associated virus (AAV)
X comprising dystrophin minigene operably linked to an expression control
X element. The dystrophin minigene in operable linkage with an expression
X control element, in a recombinant adeno-associated virus or retrovirus is
X useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
X dystrophy (BMD) in a mammalian subject. The present sequence is AAV
X vector plasmid construct containing human dystrophin minigenes, a muscle
X creatine kinase (MCK) promoter and a small polyA signal sequence
X
X Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;
X
X Query Match 53.0%; Score 1114; DB 6; Length 4414;
X Best Local Similarity 97.0%; Pred. No. 5.4e-302;
X Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
X
X 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGAGCCCAT 60
X |||||
X 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGAGCCCAT 1716
X
X 61 TCCCTCACAGCATTGGGAAGTCTCTGAAGACAAAGTCATTGGCAGTTTATTGATGGAGAG 120
X |||||
X 1717 TCCCTCACAGCATTGGGAAGTCTCTGAAGACAAAGTCATTGGCAGTTTATTGATGGAGAG 1776
X
X 121 TGAAGTAAACCTGGACCGGTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
X |||||
X 1777 TGAAGTAAACCTGGACCGGTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1836
X
X 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
X |||||
X 1837 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 1896
X
X 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
X |||||
X 1897 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 1956
X
X 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
X |||||
X 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 2016
X

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
|||
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
|||
Db 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
|||
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 540
|||
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 2196
|||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAACAACATAAGGTGCTTCA 600
|||
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAACAACATAAGGTGCTTCA 2256
|||
QY 601 AGAAGATCTAGAAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTAGT 660
|||
Db 2257 AGAAGATCTAGAAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTAGT 2316
|||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAAGAAACAACTTAAAGTATTGGG 720
|||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAAGAAACAACTTAAAGTATTGGG 2376
|||
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
|||
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 2436
|||
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGGCTTTTGTAGTCATGGCTTTTCTAGA 840
|||
Db 2437 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGGCTTTTGTAGTCATGGCTTTTCTAGA 2496
|||
QY 841 AAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
|||
Db 2497 AAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 2556
|||
QY 901 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
|||
Db 2557 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2616
|||
QY 961 GGGCAAACTGTATTCTACTCAACAAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
|||
Db 2617 GGGCAAACTGTATTCTACTCAACAAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
|||
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
|||
Db 2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2736
|||
QY 1081 ACTTGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCTCCCTGGACCTGGA 1140
|||
Db 2737 ACTTGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCTCCCTGGACCTGGA 2796
|||
QY 1141 AAAGTTTCTGCTGGCTTACAGAAGCTGA 1170
|||
Db 2797 TGAGCTGGACCTCAAGCTGCGCCAAGCTGA 2826
|||
RESULT 13
ABK81997
ID ABK81997 standard; DNA; 5417 BP.
XX
AC ABK81997;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX

WO200229056-A2.
11-APR-2002.
04-OCT-2001; 2001WO-US031126.
06-OCT-2000; 2000US-0238848P.
(UNMI) UNIV MICHIGAN.
Chamberlain JS, Harper SQ;
WPI; 2002-435334/46.
A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.
Disclosure; Fig 12; 145pp; English.
The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention
Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;
Query Match 52.5%; Score 1103.4; DB 6; Length 5417;
Best Local Similarity 99.9%; Pred. No. 5.8e-299;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 60
b 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 1158
Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
Y 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278
Y 181 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
b 1279 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
Y 241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
b 1339 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1398
Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATATTCAGAAGATGAAGA 360
b 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATATTCAGAAGATGAAGA 1458
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
b 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
Y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
b 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
Y 481 GAAACAGTTGAATGACTGCTTAACAAAAACAAGAAAGAAACAAGAAAAATGGAGGAAGA 540
b 1579 GAAACAGTTGAATGACTGCTTAACAAAAACAAGAAAGAAACAAGAAAAATGGAGGAAGA 1638

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
Db 1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTGGGAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTGGGAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
QY 841 AAAAGAGATGCAGTGAACAAAGATTCAACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAGATGCAGTGAACAAAGATTCAACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAAGATCTTTCTTCAACACTGAAGAATAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAAGATCTTTCTTCAACACTGAAGAATAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGGAAGCATGGTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAAGACGGGAAGCATGGTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGA AAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGA AAAAGAGTACAGCACAGACT 2203

RESULT 14
ABK82005
ID ABK82005 standard; DNA; 11241 BP.
XX
AC ABK82005;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin, full length HDMD.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
OS Homo sapiens.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.

XX Example 2; Fig 23; 145pp; English.
PS The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a human dystrophin polynucleotide
CC sequence used in the creation of the mini-dystrophin peptides of the
CC invention
XX
SQ Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 U; 0 Other;

Query Match 52.5%; Score 1103.4; DB 6; Length 11241;
Best Local Similarity 99.9%; Pred. No. 8.7e-299;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGAGCCCAT 60
Db |||
1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGAGCCCAT 1158

2Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATGTGAGAG 120
Db |||
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATGTGAGAG 1218

2Y 121 TGAAGTAACTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db |||
1219 TGAAGTAACTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

2Y 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||
1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

2Y 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTGG 300
Db |||
1339 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTGG 1398

2Y 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db |||
1399 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458

2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db |||
1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518

2Y 421 TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATFCAGAAACT 480
Db |||
1519 TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATFCAGAAACT 1578

2Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db |||
1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638

2Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAACTACAAACAATAGGTGCTTCA 600
Db |||
1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAACTACAAACAATAGGTGCTTCA 1698

2Y 601 AGAAGATCTAGAAACAAGACAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db |||
1699 AGAAGATCTAGAAACAAGACAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 1758

2Y 661 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAGAACAACACTTAAGTATTGGG 720
Db |||
1759 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAGAACAACACTTAAGTATTGGG 1818

2Y 721 AGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAAGACAT 780
Db |||
1819 AGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAAGACAT 1878

Qy 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCGATGGCTTTCAGA 840
Db |||
1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCGATGGCTTTCAGA 1938

Qy 841 AAAAGAAGATGCGAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
Db |||
1939 AAAAGAAGATGCGAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db |||
1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058

Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db |||
2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy 1021 CCAGAAGACCGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db |||
2119 CCAGAAGACCGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178

Qy 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db |||
2179 ACTTGAAAAGAGTACAGCACAGACT 2203

RESULT 15
ABK82002
ID ABK82002 standard; DNA; 11443 BP.
XX
AC ABK82002;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Disclosure; Fig 17; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;

Query Match		52.5%;	Score 1103.4;	DB 6;	Length 11443;
Best Local Similarity		99.9%;	Pred. No. 8.8e-299;		
Matches 1104;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
y	1	GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGGAGCCCAATT	60		
b	1099	GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGGAGCCCAATT	1158		
y	61	TCCTTCACAGCATTTGGAAGCTCCTGAAACACAAGTCATTTGGCAGTTCATTGTATGGAGAG	120		
b	1159	TCCTTCACAGCATTTGGAAGCTCCTGAAACACAAGTCATTTGGCAGTTCATTGTATGGAGAG	1218		
y	121	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTTC	180		
b	1219	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTTC	1278		
y	181	TGCTGAGGACACATTGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGGTCAAAAGA	240		
b	1279	TGCTGAGGACACATTGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGGTCAAAAGA	1338		
y	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	300		
b	1339	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	1398		
y	301	TAATATTCTACAANTGGGAAGTAAGCTCAATGGAACAGGAAAAATTATCAGAAGATGAAGA	360		
b	1399	TAATATTCTACAANTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTATCAGAAGATGAAGA	1458		
y	361	AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC	420		
b	1459	AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC	1518		
y	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAACT	480		
b	1519	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAACT	1578		
y	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA	540		
b	1579	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA	1638		
y	541	GCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA	600		
b	1639	GCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA	1698		
y	601	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT	660		
b	1699	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT	1758		
y	661	TGATGAATCTATGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG	720		
b	1759	TGATGAATCTATGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG	1818		
y	721	AGATCGATGGGCAAAACATCTGTAGTGGACAGAAGACCGCTGGGTCTTTTACAGACAT	780		
b	1819	AGATCGATGGGCAAAACATCTGTAGTGGACAGAAGACCGCTGGGTCTTTTACAGACAT	1878		
y	781	CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTGCA TGGCTTTCAGA	840		
b	1879	CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTGCA TGGCTTTCAGA	1938		
y	841	AAAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAAATGAATGTT	900		
b	1939	AAAAAGAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAAATGAATGTT	1998		
y	901	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	960		
b	1999	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	2058		
y	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1020		
b	2059	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	2118		

QY	1021	CCAGAAAGACGGAAGCATGGCTGGATAA	1080
Db	2119	CCAGAAAGACGGAAGCATGGCTGGATAA	2178
QY	1081	ACTTGAAAAAGAGTACACACAGACT	1105
Db	2179	ACTTGAAAAAGAGTACACACAGACT	2203

Search completed: April 4, 2004, 18:19:26
Job time : 579.141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model
un on: April 4, 2004, 12:04:07 ; Search time 5477.46 Seconds
(without alignments)
16625.161 Million cell updates/sec
itle: US-09-845-416-2_COPY_900_3000
erfect score: 2101
equence: 1 gagctatgcctacacacagg.....acctcagcaactctggaagac 2101
coring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
earched: 3470272 seqs, 21671516995 residues
otal number of hits satisfying chosen parameters: 6940544
inimum DB seq length: 0
aximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
atabase : GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1293	61.5	8689	6	AX538622	AX538622 Sequence
2	1131	53.8	5952	6	AR304538	AR304538 Sequence
3	1131	53.8	5952	6	AX114289	AX114289 Sequence
4	1103.4	52.5	5417	6	AX538619	AX538619 Sequence
5	1103.4	52.5	11443	6	AX538624	AX538624 Sequence
6	1103.4	52.5	12057	6	AX538627	AX538627 Sequence
7	1103.4	52.5	13957	6	AX409637	AX409637 Sequence
8	1103.4	52.5	13957	6	AX538581	AX538581 Sequence
9	1103.4	52.5	13957	9	HUMDYS	M18533 Homo sapien
10	1103.4	52.5	13977	6	AR220819	AR220819 Sequence
11	1100.2	52.4	12446	9	HSDMDR	X14298 Human mRNA
12	997	47.5	5339	6	AX538620	AX538620 Sequence
13	996	47.4	5462	6	AX538621	AX538621 Sequence
14	929	44.2	13887	4	AF070485	AF070485 Canis fam
15	911.6	43.4	4402	6	E30220	E30220 Shortened d
16	866.6	41.2	3275	10	MUSDYS	M18025 Mouse dyster
17	866.6	41.2	13815	6	AX306153	AX306153 Sequence
18	866.6	41.2	13815	6	AX538582	AX538582 Sequence
19	866.6	41.2	13815	10	MUSDYSA	M68859 Mouse dyster
20	866.6	41.2	19307	6	AR093392	AR093392 Sequence
21	866.6	41.2	19307	6	AR142592	AR142592 Sequence
22	718.8	34.2	4075	6	E30221	E30221 Shortened d
23	714.6	34.0	4402	6	E30219	E30219 Shortened d
24	635.8	30.3	13575	5	GGDYS	X13369 Chicken mRN
25	624.2	29.7	630	9	HSDMDF1	X06179 Human fetal
26	527.8	25.1	1966	9	HSDMDA1	X06178 Human adult
27	514.6	24.5	1481	9	BC009242	BC009242 Homo sapi
28	409.2	19.5	3747	6	E30218	E30218 Shortened d
29	387	18.4	387	6	AX538612	AX538612 Sequence
30	361.4	17.2	10302	6	AX538583	AX538583 Sequence
31	361.4	17.2	10302	9	HSMUPS	X69086 H.sapiens m
32	350.4	16.7	10705	10	RNAJ2967	AJ002967 Rattus no
33	348	16.6	348	6	AX538611	AX538611 Sequence
34	344.6	16.4	1898	10	BC024140	BC024140 Mus muscu
35	344.4	16.4	3521	5	AF339031	AF339031 Danio rer
36	343	16.3	2341	10	BC062163	BC062163 Mus muscu
37	343	16.3	2394	10	BC050183	BC050183 Mus muscu
38	343	16.3	6045	6	A63605	A63605 Sequence 7
39	343	16.3	6045	6	AR281528	AR281528 Sequence
40	343	16.3	10320	6	A63607	A63607 Sequence 9
41	343	16.3	10320	6	AR281529	AR281529 Sequence
42	336.6	16.0	6059	6	AX107972	AX107972 Sequence
43	336.6	16.0	11096	6	AX538584	AX538584 Sequence
44	336.6	16.0	11096	10	MMY12229	Y12229 M.musculus
45	333	15.8	333	6	AX538589	AX538589 Sequence

ALIGNMENTS

RESULT 1	AX538622	AX538622	8689 bp	DNA	linear	PAT 23-NOV-2002
AX538622	Sequence 42 from Patent WO0229056.					
LOCUS	AX538622					
DEFINITION	AX538622					
ACCESSION	AX538622.1	GI:25271171				
VERSION						
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Chamberlain, J.S. and Harper, S.Q.					
TITLE	Mini-dystrophin nucleic acid and peptide sequences					
JOURNAL	Patent: WO 0229056-A 42 11-APR-2002;					
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)					

FEATURES		Location/Qualifiers	
source	1.	.8689	
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Synthetic"		
ORIGIN			
Query Match	61.5%;	Score 1293;	DB 6; Length 8689;
Best Local Similarity	72.5%;	Pred. No. 3.8e-296;	
Matches 2101;	Conservative 0;	Mismatches 0;	Indels 798; Gaps 1;
QY	1	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	60
Db	1099	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	1158
QY	61	TCCTTCACAGCATTTGGAAGCTCCTGGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG	120
Db	1159	TCCTTCACAGCATTTGGAAGCTCCTGGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG	1218
QY	121	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC	180
Db	1219	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC	1278
QY	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240
Db	1279	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338
QY	241	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	300
Db	1339	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	1398
QY	301	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA	360
Db	1399	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA	1458
QY	361	AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	420
Db	1459	AACTGAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	1518
QY	421	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT	480
Db	1519	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT	1578
QY	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA	540
Db	1579	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA	1638
QY	541	GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAACATAAGGTGCTTCA	600
Db	1639	GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAACATAAGGTGCTTCA	1698
QY	601	AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTAGT	660
Db	1699	AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTAGT	1758
QY	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG	720
Db	1759	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG	1818
QY	721	AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTTCCTTTTACAAGACAT	780
Db	1819	AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTTCCTTTTACAAGACAT	1878
QY	781	CCTTCTCAAATGGCAACGCTTACTGAGAACAGTGCCTTTTATGTCATGGCTTTCAGA	840
Db	1879	CCTTCTCAAATGGCAACGCTTACTGAGAACAGTGCCTTTTATGTCATGGCTTTCAGA	1938
QY	841	AAAAGAAAGATGCAGTGAACAAGATTTCACAAACCTGGCTTTTAAAGATCAAAATGAAATGTT	900
Db	1939	AAAAGAAAGATGCAGTGAACAAGATTTCACAAACCTGGCTTTTAAAGATCAAAATGAAATGTT	1998
QY	901	ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	960

Db	1999	ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT	2058
QY	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1020
Db	2059	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	2118
QY	1021	CCAGAAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	1080
Db	2119	CCAGAAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	2178
QY	1081	ACTTGAAAAGAGTACAGCACAG	1102
Db	2179	ACTTGAAAAGAGTACAGCACAGATTTTCAACAGACGCTGACCTAGACTCCTGGACTGACCAC	2238
QY	1103		1102
Db	2239	TATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACACACCTGTGTTACTAAGGA	2298
QY	1103		1102
Db	2299	AACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGTATGTTGGAGGTACCTGCTCTGGC	2358
QY	1103		1102
Db	2359	AGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTCTCTGCTGTGATCAAGTTAT	2418
QY	1103		1102
Db	2419	AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA	2478
QY	1103		1102
Db	2479	GAAAGCAACAAATGCAGGATTTTGAACACAGAGGCGTCCCAGATTGGAAGAACTCATTACCGC	2538
QY	1103		1102
Db	2539	TGCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGGCTAGAACAACTCATTACGGATCG	2598
QY	1103		1102
Db	2599	AATTGAAAGAAATTCAGAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCCGAGGCA	2658
QY	1103		1102
Db	2659	ACAGTTGAATGAAATGTTAAAGGATTTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGA	2718
QY	1103		1102
Db	2719	GCAGGTCTTAGACAGCCAGAGCCAAAGCTTGAGTCATGGAAGGAGGTCCTTATACAGT	2778
QY	1103		1102
Db	2779	AGATGCAATCCAAAAGAAAATCACAGAAACCAAGCAGTTGGCCCAAAGACCTCCGCCAGTG	2838
QY	1103		1102
Db	2839	GCAGACAAATGTAGATGTGGCAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGC	2898
QY	1103		1102
Db	2899	AGATGATACCAGAAAAGTCCACATGATACAGAGAAATATCAATGCCCTCTTGGAGAAGCAT	2958
QY	1103		1122
Db	2959	TCATAAAAGGGTGATGAGCGAGAGGCTGCTTTTGAAGAAACTCATAGATTACTGCAACA	3018
QY	1123	GTTCCCTGAGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAA	1182
Db	3019	GTTCCCTGAGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAA	3078
QY	1183	TGCTCTACAGGATGCTACCCGTTAAGGAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAGA	1242

db 3079 TGTCCTACAGGATGCTACCCGTAAGGAAAGCGTCTCTAGAGACTCCACAGGAGTAAAGA 3138
y 1243 GCTGATGAAACAATGSCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCA 1302
db 3139 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCA 3198
y 1303 CAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1362
db 3199 CAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 3258
y 1363 CCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTC 1422
db 3259 CCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTC 3318
y 1423 TCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTC 1482
db 3319 TCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTC 3378
y 1483 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACC 1542
db 3379 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACC 3438
y 1543 TATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGATGATAGAGGCTTCAAGAG 1602
db 3439 TATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGATGATAGAGGCTTCAAGAG 3498
y 1603 GGAATTGAAAACTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCT 1662
db 3499 GGAATTGAAAACTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCT 3558
y 1663 GACAGAGCAGCCTTTTGGAAAGACTAGAGAACTCTACAGGAGCCACAGAGAGTGCCTCC 1722
db 3559 GACAGAGCAGCCTTTTGGAAAGACTAGAGAACTCTACAGGAGCCACAGAGAGTGCCTCC 3618
y 1723 TGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 1782
db 3619 TGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 3678
y 1783 TGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATPAGATGAGACCC 1842
db 3679 TGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATPAGATGAGACCC 3738
y 1843 TGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC 1902
db 3739 TGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC 3798
y 1903 TGAGTGTATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGA 1962
db 3799 TGAGTGTATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGA 3858
y 1963 TCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAG 2022
db 3859 TCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAG 3918
y 2023 CCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAA 2082
db 3919 CCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAA 3978
y 2083 CCTCAGCACTCTGGAAGAC 2101
db 3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 2
R304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. 5952
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 53.8%; Score 1131; DB 6; Length 5952;
Best Local Similarity 68.6%; Pred. No. 1.2e-257;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGACCCCTACACGGAGCCCAAT 60
Db 891 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGACCCCTACACGGAGCCCAAT 950
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACNAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACNAAGTCATTTGGCAGTTTCATTGATGGAGAG 1010
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1070
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1071 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1131 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1190
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1250
QY 361 AACTGAAGTACAAGAGCAGATGAAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAGAGCAGATGAAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1310
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1311 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1370
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 540
Db 1371 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 1430
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAAGTACAAACAATAGGTGCTTCA 600
Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAAGTACAAACAATAGGTGCTTCA 1490
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGGTAGT 660
Db 1491 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGGTAGT 1550
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGG 720
Db 1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGG 1610
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1611 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 1670
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTTCAGA 840
Db 1671 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTTCAGA 1730
QY 841 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 900

Db 1731 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790

Qy 901 ATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAGGAAAAAGCAATCCAT 960

Db 1791 ATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAGGAAAAAGCAATCCAT 1850

Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020

Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1910

Qy 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080

Db 1911 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1970

Qy 1081 ACTTGAAGAAGTACAGCACAG----- 1102

Db 1971 ACTTGAAGAAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGT 2030

Qy 1103 ----- 1102

Db 2031 GGAAGAGATTTTGTCTAAAGGCGAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGT 2090

Qy 1103 ----- 1102

Db 2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGAAGGCGGTAAACCGTTTACTTCAAGA 2150

Qy 1103 ----- 1102

Db 2151 GCTGAGGGCAAGCAGCCTGACCTAGTCTCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210

Qy 1103 ----- 1102

Db 2211 TCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACT 2270

Qy 1103 ----- 1102

Db 2271 AGAAATGCCATCTTCTTGATGTTGGAGGTACCTGCTCTGCAGATTTCAACCGGGCTTG 2330

Qy 1103 ----- 1102

Db 2331 GACAGAACTTACCGACTGGCTTTCTCTGTCTTGATCAAGTTATAAAATCACAGAGGTGAT 2390

Qy 1103 ----- 1102

Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGA 2450

Qy 1103 ----- 1102

Db 2451 TTTGGAACAGAGCGCTCCCAGTTTGAAGAACTCATTACCGCTGCCAAAATTGAAAAA 2510

Qy 1103 ----- 1102

Db 2511 CAAGACAGCAATCAAGAGGCTAGAACAAATCATACGGATCGAATTGAAAGAATTCAGAA 2570

Qy 1103 ----- 1102

Db 2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630

Qy 1103 ----- 1102

Db 2631 AAAGGATTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTAGGACAGGC 2690

Qy 1103 ----- 1102

Db 2691 CAGAGCCAAGTTGAGTCATGGAAGGAGGGTCCCTATACAGTAGATGCAATCCAAAAAGAA 2750

Qy 1103 ----- 1102

Db 2751 AATCACAGAAACCAAGCAGTTGGCCAAGACCTCCCGCAGTGGCAGACAAATGTAGATGT 2810

Qy 1103 ----- 1102

Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTCGAGATGATACCAGAAAAGT 2870

Qy 1103 ----- 1102

Db 2871 CCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAAGCATTATATAAAGGGTGAGTGA 2930

Qy 1103 ----- ACTCATAGATTACTGCACAAGTTCCTCCCTCCCTGGACCTGGA 1140

Db 2931 GCGAGAGGCTGCTTTTGGAAAGAACTCATAGATTACTGCACAAGTTCCTCCCTGGACCTGGA 2990

Qy 1141 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200

Db 2991 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTAC 3050

Qy 1201 CCGTAAGGAAAGGCTCTCTAGAAAGCTTCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260

Db 3051 CCGTAAGGAAAGGCTCTCTAGAAAGCTTCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 3110

Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAAACAG 1320

Db 3111 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAAACAG 3170

Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTTCACAAAGACGTTT 1380

Db 3171 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTTCACAAAGACGTTT 3230

Qy 1381 GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440

Db 3231 GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAAGTCTCTCAACATTAGGTCCCA 3290

Qy 1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACITTCGTGT 1500

Db 3291 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACITTCGTGT 3350

Qy 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCC 1560

Db 3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCC 3410

Qy 1561 AGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620

Db 3411 AGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3470

Qy 1621 ACCTGTAATCATGATGATCTTTGAGACTGTACGAATATTTCTGCACAGAGCAGCCTTTTGA 1680

Db 3471 ACCTGTAATCATGATGATCTTTGAGACTGTACGAATATTTCTGCACAGAGCAGCCTTTTGA 3530

Qy 1681 AGGACTAGAGAAACTCTACCAGAGGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCCCAGAA 1740

Db 3531 AGGACTAGAGAAACTCTACCAGAGGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCCCAGAA 3590

Qy 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800

Db 3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 3650

Qy 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860

Db 3651 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 3710

Qy 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920

Db 3711 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 3770

Qy 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 1980

Db 3771 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 3830

Qy 1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040

Db 3831 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3890

Qy 2041 TCGCCAGCTTACCACTTTGGGCAATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100

Db 3891 TCGCCAGCTTACCACTTTGGGCAATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3950

2101 C 2101
3951 C 3951
AX114289
Sequence 1 from Patent WO0129243.
AX114289
AX114289.1 GI:14031259
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Paul, X.L. and Xiao, X.
Method and vector for producing and transferring trans -spliced
peptides
Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
Location/Qualifiers
1. .5952
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"
RIGIN
Query Match 53.8%; Score 1131; DB 6; Length 5952;
Best Local Similarity 68.6%; Pred. No. 1.2e-257;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;
1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGAGCCCAT 60
891 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGAGCCCAT 950
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1010
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1070
181 TGCTGAGGACACATTGCAAGCAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
1071 TGCTGAGGACACATTGCAAGCAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1130
241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
1131 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1190
301 TAATATTCTACAATTGGGAAGTAAGTGAATTTGGAACAGGAAAATTATCAGAAGATGAAGA 360
1191 TAATATTCTACAATTGGGAAGTAAGTGAATTTGGAACAGGAAAATTATCAGAAGATGAAGA 1250
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
1251 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
1311 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1370
481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 540
1371 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 1430

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 600
DB 1431 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1490
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGCTCAATTCTCTCACTCACATGGTGGTGTAGT 660
DB 1491 AGAAGATCTAGAAACAAGAACAAAGTCAGGCTCAATTCTCTCACTCACATGGTGGTGTAGT 1550
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
DB 1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610
QY 721 AGATCGATGGGCAACACATCTGTAGATGAGACAGAACCCGCTGGTCTTTTACAAGACAT 780
DB 1611 AGATCGATGGGCAACACATCTGTAGATGAGACAGAACCCGCTGGTCTTTTACAAGACAT 1670
QY 781 CCTTCTCAAATGSCAACGCTTTACTGTAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 840
DB 1671 CCTTCTCAAATGSCAACGCTTTACTGTAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 1730
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 1731 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1790
QY 901 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1791 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1850
QY 961 GGGCAAACTGTATTCACTCAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB 1851 GGGCAAACTGTATTCACTCAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1910
QY 1021 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTAGTCCAAAA 1080
DB 1911 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTAGTCCAAAA 1970
QY 1081 ACTTGAAAAGAGTACAGCACAG----- 1102
DB 1971 ACTTGAAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCCGGATGT 2030
QY 1103 ----- 1102
DB 2031 GGAAGAGATTTTGTCTAAAGGGCAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGT 2090
QY 1103 ----- 1102
DB 2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGAAGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 ----- 1102
DB 2151 GCTGAGGGCAAGCAGCCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210
QY 1103 ----- 1102
DB 2211 TCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGSAAACTGCCATCTCCAAACT 2270
QY 1103 ----- 1102
DB 2271 AGAAATGCCATCTTCTCTTGATGTTGGAGGTACCTGCTCTGSCAGATTTCACCCGGGCTTG 2330
QY 1103 ----- 1102
DB 2331 GACAGAACTTACCGACTGGCTTTCTCTGTGTGATCAAGTTATATAAATCACAGAGGGTGAT 2390
QY 1103 ----- 1102
DB 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGA 2450
QY 1103 ----- 1102
DB 2451 TTTGGAACAGAGCGCTCCCAGGTTGGAGAACTCATACCCTGCTCCCAAAATTTGAAAAA 2510

1103 ----- 1102
2511 CAAGACCAGCAATCAAGAGGCTAGAACAAATCATTCGGATCGAATTGAAAGAATTCAGAA 2570
1103 ----- 1102
2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630
1103 ----- 1102
2631 AAAGGATTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGCTCTTAGGACAGGC 2690
1103 ----- 1102
2691 CAGAGCCAAGCTTGAGTCATGGAAGGAGGCTCCCTATACAGTAGATGCAATCCAAAAGAA 2750
1103 ----- 1102
2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGT 2810
1103 ----- 1102
2811 GGCAAAATGACTTGGGCCCTGAAAACCTTCTCCGGGATTATTCTGCAGATGATACCAGAAAAGT 2870
1103 ----- 1102
2871 CCACATGATAACAGAGAAATATCAATGCCTCTTGGAGAAAGCATTCATATAAAGGGTGATGA 2930
1103 ----- ACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 1140
2931 GCGAGAGGCTGCTTTGGAGAAAACCTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 2990
1141 AAAGTTTCTTGCTTGCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTTAC 1200
2991 AAAGTTTCTTGCTTGCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTTAC 3050
1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
3051 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 3110
1261 AGACCTCCAAGGTGAAATTCGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAG 1320
3111 AGACCTCCAAGGTGAAATTCGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAG 3170
1321 CCAAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT 1380
3171 CCAAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT 3230
1381 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
3231 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3290
1441 TTTTGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 1500
3291 TTTTGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGT 3350
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3410
1561 AGCAGTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGGAATTGAAAACTAAAGA 1620
3411 AGCAGTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGGAATTGAAAACTAAAGA 3470
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCCTTTTGA 1680
3471 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCCTTTTGA 3530
1681 AGGACTAGAGAAACTCTACAGGAGCCCAAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGAA 1740
3531 AGGACTAGAGAAACTCTACAGGAGCCCAAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGAA 3590
1741 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800

Db 3591 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAA 3650
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 3651 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3710
QY 1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
Db 3711 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3770
QY 1921 CTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 1980
Db 3771 CTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 3830
QY 1981 GGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
Db 3831 GGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3890
QY 2041 TCGCCAGCTTACCACCTTTGGGCACTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAGA 2100
Db 3891 TCGCCAGCTTACCACCTTTGGGCACTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAGA 3950
QY 2101 C 2101
Db 3951 C 3951
RESULT 4
AX538619
LOCUS AX538619 5417 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 artificial sequences.
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 39 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1 .5417
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"
ORIGIN
Query Match 52.5%; Score 1103.4; DB 6; Length 5417;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCCAT 60
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCCAT 1158
QY 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTGG 300

```
1339 CCAGTTTCATCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
301 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
1399 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 420
1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 1518
421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 480
1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 1578
481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1638
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 600
1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 1698
601 AGAAGATCTAGAAACAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTCACTCAGATGGTGGTAGT 1758
1699 AGAAGATCTAGAAACAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTCACTCAGATGGTGGTAGT 1818
661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAAGGTATTGGG 720
1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAAGGTATTGGG 1818
721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTCACTCAGATGGTGGTAGT 780
1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTCACTCAGATGGTGGTAGT 1878
781 CCTTCTCAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
1879 CCTTCTCAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
841 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 900
1939 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 1998
901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
961 GGGCAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAAATTAAGTCAGTGAC 1020
2059 GGGCAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAAATTAAGTCAGTGAC 2118
1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
1081 ACTTGAAAAGAGTACAGCACAGACT 1105
2179 ACTTGAAAAGAGTACAGCACAGATT 2203
```

```
35ULT 5
538624
DCUS AX538624 11443 bp DNA linear PAT 23-NOV-2002
3FINITION Sequence 44 from Patent WO229056.
3CESSION AX538624
3RSION AX538624.1 GI:25271175
3YWORDS synthetic construct
3URCE synthetic construct
ORGANISM artificial sequences.
3PERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
```

```
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
Location/Qualifiers
source 1. .11443
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 52.5%; Score 1103.4; DB 6; Length 11443;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCTTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAATT 1158
QY 61 TCCTTACACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGSCAGTTCAATGATGGAGAG 120
Db 1159 TCCTTACACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGSCAGTTCAATGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAAACAAAGCAACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAAACAAAGCAACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTTACAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTCTTACAGACAT 1878
QY 781 CCTTCTCAATGGCAACCGTCTTACTGAAGAAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 1879 CCTTCTCAATGGCAACCGTCTTACTGAAGAAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
QY 841 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 900
Db 1939 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 1998
```


QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
DB 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
DB 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 6
AX538627
LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .12057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 12057;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGTACCCCTACAGGAGCCCAT 60
DB 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGTACCCCTACAGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAAGTATTATCGTGGCTTCTTC 180
DB 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAAGTATTATCGTGGCTTCTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1279 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGGTTGG 300
DB 1339 CCAGTTTCATACACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
DB 1399 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGSAAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480

DB 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
DB 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACAAATAGGTGCTTCA 600
DB 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACAAATAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGGTAGT 660
DB 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAGACAT 780
DB 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAGACAT 1878
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTTCAGA 840
DB 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTTCAGA 1938
QY 841 AAAAGAAGATGCAGTGAACAAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 1939 AAAAGAAGATGCAGTGAACAAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
DB 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
DB 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 7
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1. .13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"

ORIGIN
Query Match 52.5%; Score 1103.4; DB 6; Length 13957;

Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGCACCTACACGGAGCCCAATT 60
|||||
b 1099 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGCACCTACACGGAGCCCAATT 1158
|||||

Y 61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGAGAG 120
|||||
b 1159 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGAGAG 1218
|||||

Y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTC 180
|||||
b 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTC 1278
|||||

Y 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
b 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
|||||

Y 241 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
b 1339 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1398
|||||

Y 301 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
|||||
b 1399 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1458
|||||

Y 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
b 1459 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
|||||

Y 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 480
|||||
b 1519 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 1578
|||||

Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 540
|||||
b 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 1638
|||||

Y 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACATAGGTGCTTCA 600
|||||
b 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACATAGGTGCTTCA 1698
|||||

Y 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTAGT 660
|||||
b 1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTAGT 1758
|||||

Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAAGTATGGG 720
|||||
b 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAAGTATGGG 1818
|||||

Y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAAGACAT 780
|||||
b 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAAGACAT 1878
|||||

Y 781 CTTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 840
|||||
b 1879 CTTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 1938
|||||

Y 841 AAAAGAAGATGAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
|||||
b 1939 AAAAGAAGATGAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
|||||

Y 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
|||||
b 1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058
|||||

Y 961 GGGCAAACTGATTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
|||||
b 2059 GGGCAAACTGATTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
|||||

Y 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTTTGGGATAAATTTAGTCCAAAA 1080
|||||

Db 2119 CCAGNACAGCGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178

QY 1081 ACITGAAAAAGAGTACAGCACAGACT 1105
|||||

Db 2179 ACITGAAAAAGAGTACAGCACAGATT 2203
|||||

RESULT 8
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1 .13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 60
|||||
Db 1099 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 1158
|||||

QY 61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
|||||
Db 1159 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
|||||

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTC 180
|||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTC 1278
|||||

QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
|||||

QY 241 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
Db 1339 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
|||||

QY 301 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
|||||
Db 1399 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1458
|||||

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
|||||

QY 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 480
|||||
Db 1519 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 1578
|||||

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 540
|||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 1638
|||||

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACATAGGTGCTTCA 600
|||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACATAGGTGCTTCA 1698
|||||

EPRELPEBERAQNVRLLRKKQABEVNTEWEKLNLSADWQKIDETLERLQELQEBATD
ELDLKLROAEVIGKSWQPVGDLLIDSLQDHLKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNRWKLQVAVEDRVRLQHEAHRDFGPASQHLSTSVQGP
WERAISPNKVPYYINHETQTTQCDWHPKMTELYQSLADLNNRVFSAYRTAMKLRRLQKA
LCDDLLSLSAACDALDQHNLKQNDQPMIDILQIINCLTTIYDRLEQEHNNLVNPLCVD
MCLNLLNVYDTRGIRIRVLSPFKTGIISLCKAHLEDDKYRYLFKQVASTGFCDDQRRL
GLLLHDSIQIPROLGEVASFGGSNIEPSVRSQFQANNKPEIEAALFLDMRLEPQSM
VWLPVLHRVAAAEATAKHOAKNCIKCECPIIGFRYRSLKHFNYDQCSCFSPGRVAKGH
KXHPYMEVYCTPTTSGEDVRDFAKVLKNKFRTRYFAKHPRMGYPVQTVLEGDNMET
PVTLINEPVPDPSAPASSPOLSHDDTHSRIEHYASRLAEMENSNGSYLNSISPNSID
DEHLLIQHYCQSLNQDSPLSQPRSPAQILISLESERGELEIRILADLEENRNLAQEY
DRLKQCEHKHGLSPLSPPEMMPTSPQSPRDAELTAEAKLLRQHKGRLEARMQILEDH
NKQLESQHLRLRQLLEQPOABEAKVNGTTVSSPSTLSQRSDSSQPMLLRVVGSQTSDSM
GEEDLLSPQDTSITGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM"

RIGIN

Query Match 52.5%; Score 1103.4; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 GAGCTATGCCTACACAGGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAATT 60
b 1099 GAGCTATGCCTACACAGGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAATT 1158

Y 61 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
b 1159 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218

Y 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1278

Y 181 TGCTGAGGACACATTGGGAAGTAAGCTGATTGGAAACAGGAGAGATTTCTAATGATGTGGAAAGA 240
b 1279 TGCTGAGGACACATTGGGAAGTAAGCTGATTGGAAACAGGAGAGATTTCTAATGATGTGGAAAGA 1338

Y 241 CCAGTTTCATACACTGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGSTTGG 300
b 1339 CCAGTTTCATACACTGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGSTTGG 1398

Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
b 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458

Y 361 AACTGAAGTACAAGACAGATGAATCTCTCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 420
b 1459 AACTGAAGTACAAGACAGATGAATCTCTCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 1518

Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
b 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578

Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 540
b 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 1638

Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAGGTGCTTCA 600
b 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAGGTGCTTCA 1698

Y 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTACTCACATGGTGGTAGT 660
b 1699 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTACTCACATGGTGGTAGT 1758

Y 661 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG 720
b 1759 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG 1818

Y 721 AGATCGATGGGCAAAACATCTGTAGATGCACAGAACCGCTGGGTTCTTTTACAAGACAT 780
b 1819 AGATCGATGGGCAAAACATCTGTAGATGCACAGAACCGCTGGGTTCTTTTACAAGACAT 1878

Y 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTCTAGTGCAATGGCTTTCAGA 840

Db 1879 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 1938
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACCTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACCTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAAATTTAGTCCAAA 1080
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAAATTTAGTCCAAA 2178
QY 1081 ACTTGAAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAAGAGTACAGCACAGACT 2203

RESULT 10
AR220819
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source 1..13977
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 4.1e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACAGGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCTACACAGGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAATT 1158

QY 61 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCAATTGGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 1338

QY 241 CCAGTTTCATACACTGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTGG 300
Db 1339 CCAGTTTCATACACTGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTGG 1398

QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458

QY 361 AACTGAAGTACAAGACAGATGAATCTCTCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 420

```
|||||
1459 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGTTAGC 1518
2Y   |||||
421  TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAACT 480
2Y   |||||
1519 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAACT 1578
2Y   |||||
481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGGAAATGGAGGAAGA 540
2Y   |||||
1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGGAAATGGAGGAAGA 1638
2Y   |||||
541  GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 600
2Y   |||||
1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 1698
2Y   |||||
601  AGAAGATCTAGAACAAAGAACAACTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
2Y   |||||
1699 AGAAGATCTAGAACAAAGAACAACTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1758
2Y   |||||
661  TGATGAATCTAGTGGAGATCACCGAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
2Y   |||||
1759 TGATGAATCTAGTGGAGATCACCGAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
2Y   |||||
721  AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 780
2Y   |||||
1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 1878
2Y   |||||
781  CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 840
2Y   |||||
1879 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 1938
2Y   |||||
841  AAAAGAGATGCAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAAATGAAATGTT 900
2Y   |||||
1939 AAAAGAGATGCAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAAATGAAATGTT 1998
2Y   |||||
901  ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
2Y   |||||
1999 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
2Y   |||||
961  GGGCAAACTGTATTCACTCAAAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1020
2Y   |||||
2059 GGGCAAACTGTATTCACTCAAAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 2118
2Y   |||||
1021 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
2Y   |||||
2119 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
2Y   |||||
1081 ACTTGAAAAGAGTACAGCACAGACT 1105
2Y   |||||
2179 ACTTGAAAAGAGTACAGCACAGATT 2203
```

```
RESULT 11
ISDMDR      HSDMDR      12446 bp      mRNA      linear      PRI 12-SEP-1993
LOCUS       Human mRNA for dystrophin.
DEFINITION  X14298
VERSION     X14298.1 GI:30845
KEYWORDS   Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 12446)
AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
            Davies,K.E.
TITLE      Two human cDNA molecules coding for the Duchenne muscular dystrophy
            (DMD) locus are highly homologous
JOURNAL    Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE    89345106
PUBMED     2668885
REFERENCE  2 (bases 1 to 12446)
AUTHORS   Rosenthal,A.
```

```
TITLE      Direct Submission
JOURNAL    Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
            der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
            Str.10, 1115 Berlin Buch, DDR
COMMENT    see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
            M18533 cDNA were located at xl4298 pos. 496, 1772, 1965, 2449,
            3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES   Location/Qualifiers
            1..12446
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /chromosome="X chromosomal, Xp21."
             /tissue_type="muscle"
             /dev_stage="fetal and adult."
            99..11156
             /note="unnamed protein product; dystrophin (AA 1 - 3685)"
             /codon_start=1
             /protein_id="CAA32479.1"
             /db_xref="GI:30846"
             /db_xref="GOA:P11532"
             /db_xref="SWISS-PROT:P11532"
             /translation="MLWEEVEEDCYEREDVQKTFTKWNAQFSKFGKQKHENLFSDL
            ODGRRLLDLLEGLTGOKLPKEKSTRVHALNNVNKALRVLQNNNDLVNIGSTDIDVG
            NKKLTGLIWNIIILHWQVKNMKNIMAGLOPTNSEKILLSWVRQSTRNYPQVNVINFT
            TSWDGLALNALIHSHRPDLFDWNSVVCOSATQRLEHAFNIARIQOLGIEKLLDPEDV
            DTYPDKKSILMYITSLFQVLPQQVSIEAIQEVMLPRPPKVTKEEHFQLHHQMHSQ
            QITVSLAQGYERTSPKPRFKSYAYTQAAVYVTSPTSPFPSSHLEAPEDKSFSSSL
            MESEVNLDRYQTALIEEVLWSLLSAEDTLQAGEISNDVEVVKDQFHTHEGYMMDLTAH
            OGRVGNILQGLKIGTKLSEDETEVEQMNLLNSRWECLRVASWEKQSNLHRVLM
            DLQNKKLKELNDLWTKTEERTKMEEPGLDLEDLKRQQQKHVQLQEDLEQEQRVN
            SLTHMVAWVDESSEGHATAALEEQKVLGDRWANI CRWTEDRWVLQDILLKWQRLTE
            EQCLFSAWLSEKEDAVNKIHTTGFKDQDNEMSSLOKLAVLKADLEKKKQSMGKLYSIK
            QDILLSTLKNKSVTKTEAWLDFNARCWDLNVQKLEKSTAQISQAVTTQPSLTQTVM
            ETVTVTTREQILVHAQEELPPPPQKKQITVDSEIRKRLDVIDTELHSWITRSEA
            VLOSPEFAIFERKEGNSDLKEKVNAIEREKAEKFRKLQDASRSGOALVEQMVNEGVNA
            DSIKQASEQLNSRWTFECQLLSERLNWLEYQNNIIAFYNQLQQLQEQMTTAAENWLKIQ
            PTPSEPPTAIKSQLKICKDEVNRLSGLQPOIERLKIQSIAXLKEGKGPMFLDADFVAF
            TNHFQKQFSDVQAREKELQTFIDTLPPMYQETMSAIRTWVQQSETKLSIPQLSVTDY
            EIMEQRLGELQALQSSIQEQSGLYLSTTVKEMSKKAPSEISRKYQSEFEFEEIEGRWK
            KLSQQLVEHCQKLEQMNKLRKIQNHIO TLKKWMAEVDVFLKEEWPALGDS EILKKQL
            KQCELLVSDIQIOTIOPSLSNVNEGQKIKNEAPEPEFASRLETETELKELNTQWDHMCQOY
            ARKEALKGLEKTVSLQDISEMHEWMTQAE EYLDEFYKTPDEIQKAFEMKRAK
            EEAQQKEAKVKLLTESVNSVIAQAPVPAQEALEKLELTLTITNYQWLCRLNGKCKTLE
            EYWACWHELLSYLEKANKWLNVEFPLKTTENIPGGAEEISEVLDSENLMRHSNDP
            NQIRILAQTLTDGGVMDLINEELETNFNSRWRELHEEAVRRQKLLQEQSIQSAQETENS
            LHLIQESLTFIDKOLAAVIAADKVDAACMPQEAQKIQSDLTSH EISLEEMKKHNOGKEA
            AQRVLSQIDVAQKKLQDVSMKFRLFQKPAFEPQRLQESKMILDEVKMLPALETKSVE
            QEVVQSQNLHCVNLYKSLSEVKSEVMVTKTRQIVQKQTNPKPDELDERVTALKLHY
            NELGAKVTERKQQLKCLKLSRKMKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
            VAWGKATQKEIEIKQVHLKSI TEVGEALTKVLGKETVEDKLSLNSNWIATVSRAE
            EWLNLLEQYKHMETFQNVVDHITKWI IQADTLTLDSEKKKPOQKEDVLKRLKAE LND
            IRPKVDSTRDQAANLMAHGDHCRKLVEPQISELNHRFAAISHRITKTKASIP LKELE
            QPNSDIQKLLPLEABIQGVNLKEEDFNKMNEDNEGTVKELLQBGDNLQQRITDER
            KSEBIKIKQQLLTQKHNALKDLRSQRRKALEISHQWYQYKQADDLKCLDDIEKKL
            ASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSGDGAAMAVEPTQIOLSKRWRE
            IESKFAQFRRNLNFAQIHTVREETMVMTEDMPL EISYVPSTYLTETITVVSQALLEVEQ
            LLNAPDLCAKDFEDLFQBEESLKNIKDSLQOSSGRIDI IHSKKTAAALQSATPVERVKL
            QEALSQLDFQWEKVNKMYKDRQGRFDRSVBKWRRFHYDIKIFNWLTEAEQFLRKQTQI
            PENWEHAKYKYWLKELQDGIGQRQTVVRLNATGEEIIQOSSKTASILQEKLGSLNL
            RWQEVCKQLSDRKHLKEEQNILSEFQDLNEFVLWEADNIASTPLEPGKEQQLKE
            KLEQVKLLVEELPLRQILKQLNETGGPVLVSAPISPEEQDKLENKLLKQTNLQWIKVS
            RALPEKQGEIEAQIKDLQLEKKLEDEQLNHL LLLSPIRNQLIEYNQPNQEGPFD
            VKETEIAVQAKQPDVEEILSKGQHL YKEKPATQPVKRWLEDLSLSWKA VNRLLQELRA
            KQPDLAGPLTTIGASPTQTVTLVTPVTKETAISIKLEMPSSLMLEVPALADFNELRW
            ELTDWLSLDDQVKSQRMVVGDL EDINEMI IKQKATMQDLEQRRPQLEELITAAQNLK
            NKTSNQEARTIITDRIERIQNWDEVQEHQLNRRQQLNEMLKDSQWL EAKEEAEQVL
            GQARAKLESWKEGPYTVDAIQKKITETQKQAKDLRQWQTNVDVANDLAKLLRDYSAD
            DTRKVHMITENINASWRSIHKRVSEREAL EETHRLLQQFPFLDLEKFLAWL TEATEATTA
            NVLQDATRKERLLEDKSGVKMLMKQWQDLOGEIEAHTD VYHNLDENSQKILRSLEGSD
            DAVLLQRRLDNMNFKWSLRKKS LNI RSHLEASSDQWKRLHLSLQELLVWLQKDEL
            SRQAPIGGDFPAVQKQNDVHRAFKRELKTPKPIVMSTDWQKIDETLERLQELQEA TD
            EPRELPPERAQNVTRLRKQAE E VNTWEKLN LHSADWQKIDETLERLQELQEA TD
```

ELDLKLRQAEVIKGSQWPVGDLLDLSLDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQSPYNLSTLEDLNRWKLQVAVEDRVRLQHEAHRDFGPASQHFSLSTSVQGP
WERAI SPKVPYVINHETQTTCDHPKMTELYQSLADINNVRFSAVRTAMKLRRLKA
LCLDLSLAAACDADLQHNKQNDQPMIDILQIINCLTIYDRLEQEHNNLVNPLCVD
MCLNLLNVYDTRTGRIRVLNFKTGIISLCKAHLEKYRVLKQVASSSTFCDDQRRLL
GLLLHDSIQIPROLGEVASFGGNSIEPSVRSCFQFANNKPEIEAALFLDWNRLPEQSM
VWLPVLRHVAEAETAKHQAECNICKECPIIGFRYRSUKHFNVDICQSCFFSGRVAKGH
KMHYPMVEYCTPTSGEDVDRFKVLKNKFRTRFYAKHPRMGYLPVQTVLEGDNMET
PVTLINFWPVDSPAPSSPOLSHDTHSRIEYASRLAEMENSGSYLNDSTSPNESID
DEHLIYHQSLNQDSPLSQSPSAQILISLESEERLELERILADLEENRNRLQAEY
DRLKQOHEHKGSLPLSPPEMPTSPQSPRDAELIABAKLLRQHKRLEARMQILEDH
NKQLESQHLRLQLEQPAEAKVNGTTVSSPSTLQSDSSQPMLLRVVGSQTSDSM
GEEDLLSPQDTSGLLEVMQNLNNSFPSSRGRNTPGKPMREDTM"

RIGIN

Query Match		52.4%;	Score 1100.2;	DB 9;	Length 12446;
Best Local Similarity		99.7%;	Pred. No. 2.4e-250;		
Matches 1102;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
y	1	GAGCTATGCCTACACAGGGCTGCTTATGTCACCACTCTGTACCCCTACACGGAGCCCAT	60		
b	989	GAGCTATGCCTACACAGGGCTGCTTATGTCACCACTCTGTACCCCTACACGGAGCCCAT	1048		
y	61	TCCTTCACAGCATTGTGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG	120		
b	1049	TCCTTCACAGCATTGTGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG	1108		
y	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC	180		
b	1109	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC	1168		
y	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240		
b	1169	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1228		
y	241	CCAGTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGTGG	300		
b	1229	CCAGTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGTGG	1288		
y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA	360		
b	1289	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA	1348		
y	361	AACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	420		
b	1349	AACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	1408		
y	421	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAACT	480		
b	1409	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAACT	1468		
y	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA	540		
b	1469	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA	1528		
y	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA	600		
b	1529	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA	1588		
y	601	AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT	660		
b	1589	AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT	1648		
y	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATGGG	720		
b	1649	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATGGG	1708		
y	721	AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAAGACAT	780		
b	1709	AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAAGACAT	1768		
y	781	CCTTCTCAAATGGCAACCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA	840		

Db	1769	CCTGCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTTCAGA	1822
QY	841	AAAAGAAGATGCAGTGAAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT	900
Db	1829	AAAAGAAGATGCAGTGAAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT	1888
QY	901	ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	960
Db	1889	ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	1948
QY	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC	1020
Db	1949	GGGCAAACTGTATTCAATCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC	2008
QY	1021	CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	1080
Db	2009	CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA	2068
QY	1081	ACTTGAAAAGAGTACAGCACAGACT	1105
Db	2069	ACTTGAAAAGAGTACAGCACAGATT	2093
RESULT 12			
AX538620			
LOCUS	AX538620	5339 bp	DNA linear PAT 23-NOV-2002
DEFINITION	Sequence 40 from Patent WO0229056.		
ACCESSION	AX538620		
VERSION	AX538620.1 GI:25271166		
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1		
AUTHORS	Chamberlain, J.S. and Harper, S.Q.		
TITLE	Mini-dystrophin nucleic acid and peptide sequences		
JOURNAL	Patent: WO 0229056-A 40 11-APR-2002;		
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)		
FEATURES	Location/Qualifiers		
source	1..5339		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Synthetic"		
ORIGIN			
Query Match 47.5%; Score 997; DB 6; Length 5339;			
Best Local Similarity 100.0%; Pred. No. 7.8e-226;			
Matches 997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1105	TCATAGATTACTGCAACAGTTCCTGGACCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA	1164
Db	1546	TCATAGATTACTGCAACAGTTCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA	1605
QY	1165	AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA	1224
Db	1606	AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA	1665
QY	1225	CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC	1284
Db	1666	CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC	1725
QY	1285	TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGA	1344
Db	1726	TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGA	1785
QY	1345	AGGTTCCGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAG	1404
Db	1786	AGGTTCCGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAG	1845
QY	1405	TGAATTCGGAAAAAGTCTCTCAACATTAGTTCCTTTTGGAAAGCCAGTTCTGACCAGTG	1464
Db	1846	TGAATTCGGAAAAAGTCTCTCAACATTAGTTCCTTTTGGAAAGCCAGTTCTGACCAGTG	1905

QY 1465 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1524
| | | | |
Db 1906 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1965
| | | | |
QY 1525 ATTAAGCCGGCAGGCACCTATTGGAGCGACCTTTCCAGCAGTTTCCAGAGCAGAACGATGT 1584
| | | | |
Db 1966 ATTAAGCCGGCAGGCACCTATTGGAGCGACCTTTCCAGCAGTTTCCAGAGCAGAACGATGT 2025
| | | | |
QY 1585 ACATAGGGCCTTCAAGAGGGAAATTGAAAACCTTAAAGAACCTGTGAATCATGAGTACTTTGA 1644
| | | | |
Db 2026 ACATAGGGCCTTCAAGAGGGAAATTGAAAACCTTAAAGAACCTGTGAATCATGAGTACTTTGA 2085
| | | | |
QY 1645 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGCTAGAGAACTCTACCCAGGA 1704
| | | | |
Db 2086 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGCTAGAGAACTCTACCCAGGA 2145
| | | | |
QY 1705 GCCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCA 1764
| | | | |
Db 2146 GCCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCA 2205
| | | | |
QY 1765 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1824
| | | | |
Db 2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2265
| | | | |
QY 1825 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGA 1884
| | | | |
Db 2266 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGA 2325
| | | | |
QY 1885 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCT 1944
| | | | |
Db 2326 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCT 2385
| | | | |
QY 1945 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCC 2004
| | | | |
Db 2386 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCC 2445
| | | | |
QY 2005 TCTGAAAGAGAACGTGAGCCACCGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAT 2064
| | | | |
Db 2446 TCTGAAAGAGAACGTGAGCCACCGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAT 2505
| | | | |
QY 2065 TCAGCTCTACCGTATAACCTCAGCACTCTGGAAGAC 2101
| | | | |
Db 2506 TCAGCTCTACCGTATAACCTCAGCACTCTGGAAGAC 2542
| | | | |

RESULT 13
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1..5462
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN

Query Match 47.4%; Score 996; DB 6; Length 5462;
Best Local Similarity 100.0%; Pred.No. 1.3e-225;
Matches 996; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1106 CATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAA 1165
| | | | |
Db 1670 CATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAA 1729
| | | | |
QY 1166 GCTGAAACAACTGCGCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGAC 1225
| | | | |
Db 1730 GCTGAAACAACTGCGCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGAC 1789
| | | | |
QY 1226 TCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCT 1285
| | | | |
Db 1790 TCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCT 1849
| | | | |
QY 1286 CACACAGATGTTTATCAACACCTGGATGAAAAACAGCCCAAAAATCCTGAGATCCCTGGAA 1345
| | | | |
Db 1850 CACACAGATGTTTATCAACACCTGGATGAAAAACAGCCCAAAAATCCTGAGATCCCTGGAA 1909
| | | | |
QY 1346 GGTTCGGATGATGAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGT 1405
| | | | |
Db 1910 GGTTCGGATGATGAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGT 1969
| | | | |
QY 1406 GAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGATTCTGACCCAGTGG 1465
| | | | |
Db 1970 GAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGATTCTGACCCAGTGG 2029
| | | | |
QY 1466 AAGCGTCTGCACCTTTCTGCGAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 1525
| | | | |
Db 2030 AAGCGTCTGCACCTTTCTGCGAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 2089
| | | | |
QY 1526 TTAAGCCGGCAGGCACCTATTGGAGGCACTTCCAGCAGTTTCCAGAGCAGATTGAGAAACGATGTA 1585
| | | | |
Db 2090 TTAAGCCGGCAGGCACCTATTGGAGGCACTTCCAGCAGTTTCCAGAGCAGATTGAGAAACGATGTA 2149
| | | | |
QY 1586 CATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTGTAATCATGAGTACTCTTGAG 1645
| | | | |
Db 2150 CATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTGTAATCATGAGTACTCTTGAG 2209
| | | | |
QY 1646 ACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGGAG 1705
| | | | |
Db 2210 ACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGGAG 2269
| | | | |
QY 1706 CCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAG 1765
| | | | |
Db 2270 CCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAG 2329
| | | | |
QY 1766 GCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGA 1825
| | | | |
Db 2330 GCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGA 2389
| | | | |
QY 1826 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGSCCAGGATGAGCTGGAC 1885
| | | | |
Db 2390 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGSCCAGGATGAGCTGGAC 2449
| | | | |
QY 1886 CTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGSCCGTGGCGGATCTCCTC 1945
| | | | |
Db 2450 CTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGSCCGTGGCGGATCTCCTC 2509
| | | | |
QY 1946 ATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCT 2005
| | | | |
Db 2510 ATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCT 2569
| | | | |
QY 2006 CTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATT 2065
| | | | |
Db 2570 CTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATT 2629
| | | | |
QY 2066 CAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2101
| | | | |
Db 2630 CAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2665
| | | | |

RESULT 14
AF070485
LOCUS AF070485 13887 bp mRNA linear MAM 09-DEC-1998
DEFINITION Canis familiaris dystrophin mRNA, complete cds.

CESSION AF070485
ERSION AF070485.1 GI:3982750
YWORDS
OURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 13887)
AUTHORS Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
EATURES Location/Qualifiers
1..13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Breed Golden retriever"
/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="cystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEVCYEREDVQKKTFTKWNAQFSKFGKHIEHLFSDLL
QDGRRLDLLECLTGQKLPKEKGSRTVHALNNVKA.RVLQKNNVDLVNIGSTDIVDG
NHKLTGLIWNILHQQVQVMKIMAGLQOTNSEKILLSWVRQSTRNYPQVNVINFT
TSWSDGLKALNALIHSHRPDLFDWNSVVCQSATQRIEHLAFNIAKYGLIEKLLDPEDV
ATTYDPKKSILMYITSLEFQVLPQOVSIEAIOEVMELPRPSQVTRREEHQIHQWHSYQ
QITVSLAQGERAPSPFKPRFSYAYTQAAVYTTSDPTSPSPSPSQHLETPEDKSFGRS
LTETEANLDSYQTALEEVLSWLLSADALQAQGEISNDVEEVKEQFHTHEGYMMDLTS
HQVRGNVLQGLIGTGKLSDETEVEQEQMNLNWECLRVASMEKOSNLHKVL
MDLRQQLKELNDWLTKTEERTRKMEKEPLGPDIEDLKRVQQHVKVLEDELEQEVRV
NSLTHMVVVVDESJDHATAALEOLKVLGDRWANI.CRWTEDRVLLQDILLKWQFT
EQCLFSANLSEKEDAVNKIHTTGFKDQSEVLNQLKVLKTDLEKKKQTMDKLCSL
NQDLLSALTNTVAHKMEAWLDNFQORNDNLVQKLESAQISQAVTTTPSLTQTTV
METVTMTVTRHILVKAQEEELPPPPQKKRQIIVDSEIRKRLDVIDITELHSWITRSE
AVLQSPFATYRKEGNFSDLKEKVNAIEREKAERKLDQASRSQAALVEQMVNNEGVN
ADSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIITFYNLQQLQLEQMTTFAENWLKT
OPTTTSEPTATKSKQKICKDEINRLSALQPQIERLKIOSIALKEKGQGMFLDADFVA
FTNHFQNVFADVQAREKELQTI FDSLPPMRYQETMSTILTIIQQSETKLSIPOVTYTE
YIMEQRLQALQSSLOEQQGNLYLSTTVKEMSKAPLSDISRKYQSEFEEIEGR
WKLLSSQLVEHCQKLEEQMAKRLKIQNHIKTLKKWITEVDVFLKEEWPALGDSILKR
QLKQCRLLVNDIQTIQFSLNSVNEGQKWKNEAEPEFAGRLTELRNLNTQWDMCRQ
VYARKEALGGLDKTVSLQKLSNEHEWMTQAESEYEQEDIEFKYTPDELQTAVEEMKR
AKEEAQKQKAEKVKLLTESVNSVIAQAPPAQKALKKELDTLTNTYQWLCTRLNGCKT
LEEVAWCWHELLSYLEKANKWLSEVEVKLKTENISGGAEIAEVLDSLENLMQHSDE
NPNQIRILAQTLTDGGVMDLINEELETENSRWREHHEEAVRRQKLLQESIQSAQEIIE
KSLHLIQESLSIDKQLAAIADKVDAQMPQEAQRIQSDLTSHSISLEEMKKHNGQK
ETAQRVLSQIDVAQKKLQDVSMKFRFLQKPFANFEQRLQESKMILDEVKMLPALETKS
VEQEVVQSLNHCNLYKSLSEVSEVMVITKGRQIVQKKQTENPRELDERVTALKL
HYNELGAKVTERKQOLEKCLKLSRMRKEMNALTEWLAATDMELTKRSAVEGWPNSLD
SEVANGKATQKIEIKQKHLKSVTEVGEALKTVLKEKMLVEDKLSLLSNWIAVTSR
ABEWNLLLEVKHMETFDQNVDIYTNWIIQADALLDESEKKKPOQKEDILKRLKAEM
NDIRPKVDSTRQAANLMANRGDHCXKVEPKISELNHRFAAISHRIKTGKASIPLKE
LEQFNSDIQKLLPLEAEIQQGVNLKEEDFNKDMSEDNEGTVKELLQRGDNLQQRITD
ERKREEIKIKQQLIQTKNALKDLRSQRRKALEISHOWYQYKQADDDLLKCLDDIEK
KLASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSDEGAMAVEPTQILSKRW
REIESKFAQFRRLNFAQIHTVHEESVAMTEDMPLEISYVPSTYLTETIHSVQALSEV
BELLNAPDLCAQDFEDLFKQEESLKNIKDSLQOISGRIDIHNKKTAAHLSATPAERA
KLQEALSRLDFOWERVNNMYKORQGRFDRSVEKWRFRFYDMKILNQWLTEAEQFLKKT
QIPENWEHAKYKWLKELQDGIQGRQSVRVNLNATGEEIIQQSSKTDASILQEKLGSL
NLRWQEVCKQLAERKKRLEEQKNILSEFQDVNEFVLWEADNDVANIPLERQNEQQL
KEKLEQVKLLAEELPLRQGIKQLNETGGTVLVSAPLSPEEQDKLENKLLQTNLQWIK
VSRNLPEKQEEIHAHVKDLGQLEEQNLHLLWLSPIRNOLEIYNQPNQTPGPFPIKEIE
VAVQAKQPDVEGILSKGOHLYKEKPATQPAKRKLEDDSSDWKVVTQLLOELRAKQPGP
APGLTTVRAPPSQVTVLVTPAVTKETAISKLEMPSSLLLEVPALADFNRAWTELTDW
LSLLDRVIKSQRVNMGDLEDINEMIIKQKATLQDLEQRPPQLEELITAAQNKNKTSN
QEARTIITDRIERIQSQWDEVQEHQNRRLQLTEMLKDSQTQWLEAKEEAEQVLQARA
KLESWKEAPYTVDAIQKKTETKQLAKDLRQWQINVPDANDLALKLRDYSADDTRKV
HMWITENINASWASHKRLSERAALETHRLLQCFPLDEKFLAWLTEAETTANLQD
ATHKERLSDSKGVRBELMKQWQDLQGEIEAHTDIYHNLDENGKQVLRLESGSDAALL
QRRLDNMNFKWSELKKSINIRSHLEASSDQWKRLHLSLQELLVWLKLDDELSRQAP

IGDFFPAVQKQNDVHRAFKRELKTKPEVIMSTLETVRIFLTEPLEGLEKLYQEPREL
PPEERAQNVTRLLRKQAEVNTQWEKLVNHSADWQRKIDAEALERLQELQEAATDELDK
LRQAEVIKGSWQPVGDLILDSLDQHLKVKALRGEITPLKENVSYVNDLARQLTTLGI
QLSPYNLTLEDLNRWKLLOVAIEDRIQLHEAHRDFGPASQHFSLTSVQGGPWERAI
SPNKVPYIINHETQTTCCDHPKMTELYSLADLNNRVFSAYRTAMKRLRLQKALCDDL
LSLSAACDALDQHNLKQNDQPMDILQVINCLTTIYDRLEQEHNNLVNVLQVDMCLNW
LLNVYDTGRTRIRVLSFKTGIIISLCKAHLEDKRYLFLFKQVASTSTGFCDDQRRLLGLLH
DSIQIPROLGEVAFSGSNIEPSVRSFQFANNKPEIEAALFLDWMRLEPQSMVWLPV
LHRVAAAEATAKQAKNICKECPIIGFYRSLKHFNYDICQSCFSGRVAKGHMKHY
MVEYCTPTTSGEDVRDFAKVLKNKFRKRYFAKHPRMGYLPVQTVLEGDNMETPTVLI
NFWPVDSPAPASSQLSHDDTHSRIEHYASRLKKMENSNGSYLNDISIPNESIDDEHLL
IQHWYRSLNQESPLSQPRSPAQIILISLESEERGELELIRLADLEGRNRNLQAEYDRLKQ
QEHYKGLSPSPPEMPTSPQSPRDAELIAEAKLLRHQKGRLEARMQILEDHNKQLE
SQLHRLRQLLEQQAQAEAKVNGTTVSSPSTSLORSOSSQPMLLRVRVGSQTSSEMGEDL
LSPPODTSTGLEEYMEQLNHSFPSSRGENTPGKPMREDTM"
11333..>13887
11789..11828
3'UTR
repeat_region
/rpt_type=tandem
/rpt_unit="ca"

ORIGIN

Query Match 44.2%; Score 929; DB 4; Length 13887;
Best Local Similarity 90.0%; Pred. No. 1.1e-209;
Matches 995; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 60
Db 1183 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 1242
QY 61 TCCTTCACAGCATTTGGAAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGAGAGAG 120
Db 1243 TCCTTCACAGCATTTGGAAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGAGAGAG 1302
QY 121 TGAAGTAAACCTGGACCGTCTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1303 CGAAGCAAAACCTGGACAGTTCATCAAAACAGCTTTTGAAGAAGTACTCTCGTGGCTTCTTTC 1362
QY 181 TGCTGAGGACACATTTGCAAGCAAGGAGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
Db 1363 AGCTGAGGATGCACTGCAAGCCCAAGGAGAGAGATTTCCTAATGATGTGGAAGAAGTGAAGA 1422
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCGGGTGG 300
Db 1423 ACAATTCATCTCATGAGGGGTATATGATGGACITGACATCCCATCAGGGACGGGTGCG 1482
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1483 TAATGTTCTCAACTGGGAAGTCAACTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1542
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1543 AACCGAAGTGCAGGAACAATGAATCTCCTCAATCAAGATGGGAATGCCTCAGGGTAGC 1602
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 1603 TAGCATGGAACAAACAAAGCAATTTACATAAAGTTCTTAATGGATCTCCAGAAATCAGCAACT 1662
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 1663 GAAAGAGTTAAATGACTGGCTTAACAAAAACAGAAGAGAGAAACAAGGAAAAATGGAGGAAGA 1722
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACCGCCCAAGTACAACAACATAAGTGTCTTCA 600
Db 1723 GCCCCCTTGGACCTGATATTGAAGACCTTAAACCGCCCAAGTACAACAACATAAGTGTCTTCA 1782
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCCTCACTCACATGGTGGTGTAGT 660
Db 1783 AGAAGACTTAGAACACAGGAACAAAGTCAAGGTCAATTCCTCACTCATATGGTGGTGTAGT 1842
QY 661 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 720
Db 1843 CGATGAATCTAGTGGAGACCATGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTACTGGG 1902

2Y	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTCTTTACAAGACAT	780
2b	1903	AGATCGATGGGCAAAACATCTGTAGGTGGACAGAAGATCGCTGGGTCTCTTTACAAGACAT	1962
2Y	781	CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTTCAGA	840
Db	1963	CCTCCTAAAATGGCAGCGTTTTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTTCGGA	2022
2Y	841	AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAATGTT	900
2b	2023	GAAGGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAGGATCAAAAGTGAAGTGT	2082
2Y	901	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	960
Db	2083	ATCAAATCTTCAGAAACTGGCTGTCTTAAATAACAGATCTGGAAAAAGAACAGCAAAACCAT	2142
2Y	961	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1020
Db	2143	GGCAAACTCTGCTCACTCAACCAAGACCCTTCTTTTCAGCGCTGAAAAACACAGTGGTAGC	2202
2Y	1021	CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA	1080
Db	2203	CCACAAGATGGAAGCATGGCTGGACAACCTTTGCCCAGCGCTGGGATAATTTAGTCCAGAA	2262
QY	1081	ACTTGAAAAGAGTACAGCACAGACT	1105
Db	2263	ACTTGAAAAAAGTTTCAGCACAGATT	2287

```

RESULT 15
E30220
LOCUS      Shortened dystrophin.
DEFINITION
ACCESSION  E30220
VERSION    E30220.1 GI:13017027
KEYWORDS   JP 1999318467-A/3.
SOURCE     unidentified
           unidentified
           unclassified.
           1 (bases 1 to 4402)
REFERENCE  1 (bases 1 to 4402)
AUTHORS    Sinichi, T.
TITLE      Shortened dystrophin
JOURNAL    Patent: JP 1999318467-A 3 24-NOV-1999;
           SCIENCE & TECH AGENCY, NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY

COMMENT    OS Unidentified
           PN JP 1999318467-A/3
           PD 24-NOV-1999
           PF 08-MAY-1998 JP 1998142134
           PR
           PI SINICHI TAKEDA
           PC C12N15/09,A61K48/00,C12N15/00
           CC Strandedness: Both;
           CC Topology: Linear;
           FH Key
           FT source
           FT Location/Qualifiers
              1..4402
              /organism='Unidentified'.

FEATURES   Location/Qualifiers
           1..4402
           /organism="unidentified"
           /mol_type="genomic DNA"
           /db_xref="taxon:32644"
           source

```

ORIGIN

Query Match	43.4%	Score 911.6;	DB 6;	Length 4402;
Best Local Similarity	99.6%	Pred. NO. 1.5e-205;		
Matches 914; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGGCCATT	60
DB	1095	GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGGCCATT	1154
QY	61	TCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTTGGCAGTTCATTGATGGAGAG	120

Search completed: April 5, 2004, 02:39:35
Job time : 5504.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
M nucleic - nucleic search, using sw model
un on: April 4, 2004, 16:52:43 ; Search time 2799.18 Seconds
(without alignments)
17079.796 Million cell updates/sec
itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgctacacacagg.....taccagtcttagctgacct 1601

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
earched: 27513289 seqs, 14931090276 residues
otal number of hits satisfying chosen parameters: 55026578
inimum DB seq length: 0
aximum DB seq length: 20000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase :				EST:*			
1:	em_estba:*			1:	em_estba:*		
2:	em_esthum:*			2:	em_esthum:*		
3:	em_estin:*			3:	em_estin:*		
4:	em_estmu:*			4:	em_estmu:*		
5:	em_estov:*			5:	em_estov:*		
6:	em_estpl:*			6:	em_estpl:*		
7:	em_estro:*			7:	em_estro:*		
8:	em_htc:*			8:	em_htc:*		
9:	gb_est1:*			9:	gb_est1:*		
10:	gb_est2:*			10:	gb_est2:*		
11:	gb_htc:*			11:	gb_htc:*		
12:	gb_est3:*			12:	gb_est3:*		
13:	gb_est4:*			13:	gb_est4:*		
14:	gb_est5:*			14:	gb_est5:*		
15:	em_estfun:*			15:	em_estfun:*		
16:	em_estom:*			16:	em_estom:*		
17:	em_gss_hum:*			17:	em_gss_hum:*		
18:	em_gss_inv:*			18:	em_gss_inv:*		
19:	em_gss_pln:*			19:	em_gss_pln:*		
20:	em_gss_vrt:*			20:	em_gss_vrt:*		
21:	em_gss_fun:*			21:	em_gss_fun:*		
22:	em_gss_mam:*			22:	em_gss_mam:*		
23:	em_gss_mus:*			23:	em_gss_mus:*		
24:	em_gss_pro:*			24:	em_gss_pro:*		
25:	em_gss_rod:*			25:	em_gss_rod:*		
26:	em_gss_phg:*			26:	em_gss_phg:*		
27:	em_gss_vrl:*			27:	em_gss_vrl:*		
28:	gb_gss1:*			28:	gb_gss1:*		
29:	gb_gss2:*			29:	gb_gss2:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES			
				%			
result	Query	Match	Length	DB	ID	Description	
No.	Score	Match	Length	DB	ID	Description	
1	1091	68.1	3870	11	BC036103	BC036103 Homo sapi	
2	854.2	53.4	3056	11	AK044536	AK044536 Mus muscu	
3	833.8	52.1	5691	29	AY399453	AY399453 Homo sapi	
4	621	38.8	5697	29	AY399455	AY399455 Mus muscu	

	5	585.2	36.6	728	14	CB228986	CB228986 AGENCOURT
	6	502	31.4	5676	29	AY399454	AY399454 Pan trogl
C	7	497	31.0	595	14	CB177816	CB177816 is21c01.x
	8	481.2	30.1	1047	14	CB850319	CB850319 MRA-0070
	9	332.6	20.8	2334	11	BC011062	BC011062 Mus muscu
	10	322.6	20.1	9915	29	AY407022	AY407022 Homo sapi
	11	301.8	18.9	9691	29	AY407024	AY407024 Mus muscu
	12	299.8	18.7	9096	29	AY407023	AY407023 Pan trogl
	13	264.8	16.5	1122	14	CF109978	CF109978 Shultzomi
	14	255.4	16.0	772	12	BI250598	BI250598 602993659
	15	237.4	14.8	3753	11	AK081426	AK081426 Mus muscu
	16	223.8	14.0	826	14	CD656896	CD656896 AGENCOURT
	17	220.6	13.8	696	12	BJ075057	BJ075057 BJT075057
	18	216.4	13.5	784	12	BG212445	BG212445 RST32032
	19	208.6	13.0	630	9	AL855376	AL855376 AL855376
	20	205.2	12.8	851	13	BU201022	BU201022 603952191
	21	202.4	12.6	502	9	AL602076	AL602076 DKFZp3113B
	22	201	12.6	2874	29	AY408546	AY408546 Homo sapi
	23	199.4	12.5	2874	29	AY408548	AY408548 Mus muscu
	24	171	10.7	479	13	BQ304046	BQ304046 QV2-BT063
	25	169.8	10.6	835	12	BI553820	BI553820 603190772
	26	169.2	10.6	732	13	BU107880	BU107880 603109863
C	27	166.6	10.4	646	12	BI289102	BI289102 UI-R-DK0-
	28	166.2	10.4	907	13	BU122401	BU122401 603003073
	29	165.4	10.3	681	13	BU301653	BU301653 603609005
	30	158.6	9.9	615	14	CA377239	CA377239 655768 NC
	31	158.2	9.9	569	29	CE228097	CE228097 tigr-gss-
C	32	157.2	9.8	427	10	BE817979	BE817979 CM2-BN027
	33	154.6	9.7	405	14	CB811308	CB811308 AMGNNUC:S
	34	154.6	9.7	644	13	BU313510	BU313510 603540290
	35	152.2	9.5	1541	11	AK034383	AK034383 Mus muscu
	36	151.2	9.4	3051	11	BC036095	BC036095 Homo sapi
C	37	150.4	9.4	515	28	AZ780914	AZ780914 2M0018010
	38	149.6	9.3	697	13	BU456556	BU456556 603771141
	39	146.2	9.1	2874	29	AY408547	AY408547 Pan trogl
	40	144.8	9.0	434	13	BQ375536	BQ375536 PM1-TN012
	41	141	8.8	763	14	CB518960	CB518960 UI-M-GH0-
	42	136.4	8.5	466	14	CD549993	CD549993 B0305E01-
C	43	131.6	8.2	410	13	BQ349936	BQ349936 PM1-HT034
	44	131.6	8.2	778	14	CB524596	CB524596 UI-M-FY0-
	45	130.2	8.1	426	13	BQ319056	BQ319056 IL5-CT051

ALIGNMENTS

RESULT 1	BC036103	3870 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	BC036103	Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (CDNA clone IMAGE:5274415), with apparent retained intron.			
DEFINITION	BC036103	Homo sapiens (human)			
ACCESSION	BC036103	GI:23271310			
VERSION	BC036103.1	HTC			
KEYWORDS	HTC	Homo sapiens			
SOURCE	HTC	Homo sapiens			
ORGANISM	HTC	Homo sapiens			
REFERENCE	1	(bases 1 to 3870)			
AUTHORS	1	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Muliahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,			

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE
PUBMED
12477932
REFERENCE
2 (bases 1 to 3870)
AUTHORS
Strausberg,R.
TITLE
Direct Submission
JOURNAL
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 68.1%; Score 1091; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 9.7e-237;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACAGGAGCCCAT 60
DB 1043 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACAGGAGCCCAT 1102
QY 61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 120
DB 1103 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 1162
QY 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
DB 1163 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1222
QY 181 TGCTGAGACACATTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
DB 1223 TGCTGAGACACATTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1282
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGCGCGGGTTGG 300
DB 1283 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGCGCGGGTTGG 1342

QY 301 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
DB 1343 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1402
QY 361 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGC 420
DB 1403 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGC 1462
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATC-GAAACT 479
DB 1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1522
QY 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
DB 1523 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1582
QY 540 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCCAAGTACAACAACATAAGGTCTTCA 599
DB 1583 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCCAAGTACAACAACATAAGGTCTTCA 1642
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 659
DB 1643 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 1702
QY 660 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 719
DB 1703 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1762
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
DB 1763 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1822
QY 780 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 839
DB 1823 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 1882
QY 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 899
DB 1883 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 1942
QY 900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATAGAAAAAGAAAAAGCAATCCAT 959
DB 1943 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATAGAAAAAGAAAAAGCAATCCAT 2002
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
DB 2003 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2062
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1079
DB 2063 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2122
QY 1080 ACTTGAAAAAGAGTACAGCACAGA 1102
DB 2123 ACTTGAAAAAGAGTACAGCACAGA 2145

RESULT 2

AK044536
LOCUS
DEFINITION
Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930019F21 product:dystrophin, muscular dystrophy, full
insert sequence.
ACCESSION
AK044536
VERSION
AK044536.1 GI:26090404
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci,P. and Hayashizaki,Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3056)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

FEATURES
source

/db_xref="FANTOM DB:A930019F21"
/db_xref="MGI:2409854"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. .3056
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007858, evidence: BLASTN, 100%, match=999)"

ORIGIN

Query Match 53.4%; Score 854.2; DB 11; Length 3056;
Best Local Similarity 86.5%; Pred. No. 5.2e-183;
Matches 954; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
|||
Db 944 GAGTATGCCTTACACACAGGCTGCTTATGTGTCACACCTCTGATTCACACAGAGCCCTA 1003
|||
QY 61 TCCTTCACAGCATTGGAGCTCCTGAAGACAGTCACTTTGGCAGTTTCATTGATGGAGAG 120
|||
Db 1004 TCCTTCACAGCATTGGAGCTCCTGAAGACAGTCACTTTGACAGTTTCATTGATGGAGAC 1063
|||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
|||
Db 1064 GGAAGTAAATCTGGATAGTTAGTACCAAACTGCTTTAGAAAGAGTACTTTTCATGGCTTCTTTC 1123
|||
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
|||
Db 1124 TGCCGAGGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAGTGAAGA 1183
|||
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 300
|||
Db 1184 ACAGTTTCATCTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 1243
|||
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
|||
Db 1244 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAGAGGAAATTTATCAGAAGATGAAGA 1303
|||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 1304 AGCTGAAGTCAAGAACAAATGAATCTCCTAAATTTCAAGATGGGAATGTTCTCAGGGTAGC 1363
|||
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
|||
Db 1364 TAGCATGGAAAAACAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAAATCAGAAATT 1423
|||
QY 480 GAAAGAGTTGAATGATCTGGCTAACAAAAACAAGAAAGAAACAAGGAAATGGAGGAAGA 539
|||
Db 1424 AAAAGAACTAGATGATGCTGTTAAACAAAACCTGAAGAGAGAACTAAGAAATGGAGGAAGA 1483
|||
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAAGGTGCTTCA 599
|||
Db 1484 GCCTTTGGACCTGATCTTGAAGATCTAAAAATGCCAAGTACAACAATAAGGTGCTTCA 1543
|||
QY 600 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTTCTCACTCACTGATGGTGGTAGT 659
|||
Db 1544 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAAGGTCACTCACTCACTGATGGTAGTGGT 1603
|||
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 719
|||
Db 1604 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGAAGAAACAACCTTAAGGTACTGGG 1663
|||
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
|||
Db 1664 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT 1723
|||
QY 780 CCTTCTCAAATGGCAAGCTTTTACTGAAGAACAGTGCCTTTTTTAGTCATGGCTTTTTCAGA 839
|||
Db 1724 TCTTCTAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTTTAGTACATGGCTTTTTCAGA 1783
|||
QY 840 AAAAGAAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
|||

1784 AAAAGAGATGCAATGAAGAACATTGAGACAGTGGCTTTAAAGATCAAAATGAAATGAT 1843
900 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
1844 GTCAAGTCTTCAAAAATATCTACTTTTAAAAATAGATCTAGAAAAGAAAAGCCAACCAT 1903
960 GGGCAAACTGTATTCTACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
1904 GGAAAACTAAGTTCACTCAATCAAGATCTACTTTCCGCACTGAAAAATAAGTCAGTGAC 1963
1020 CCAGAACGCGAAGCATGGCTGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
1964 TCAAAAGATGGAATCTGGATGGAATACTTTGCAACAACGTTGGGACAAATTTAACCCAAAA 2023
1080 ACTTGAAGAAGAGTACAGCACAGA 1102
2024 ACTTGAAGAAGAGTTTCAGCACAAA 2046
AY399453 5691 bp DNA linear GSS 12-DEC-2003
Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY399453
AY399453.1 GI:39755442
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5691)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5691)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..5691
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>5691
/locus_tag="HCM0229"
Query Match 52.1%; Score 833.8; DB 29; Length 5691;
Best Local Similarity 89.1%; Pred. No. 2.5e-178;
Matches 983; Conservative 0; Mismatches 2; Indels 118; Gaps 3;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
294 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 353
61 TCCTTCACAGCATTGGAGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
354 TCCTTCACAGC-----TCCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 404
121 TGAAGTAAACCTGGACCGCTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 180

405 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 464
181 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
465 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 524
241 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
525 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 584
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
585 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 644
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
645 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 704
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
705 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 764
480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 539
765 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 824
540 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCCAAGTACAACAACATAAGGTGCTTCA 599
825 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCCAAGTACAACAACATAAGGTGCTTCA 884
600 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 659
885 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 944
660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 719
945 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 1004
720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
1005 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1064
780 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 839
1065 CCTTCTCAAATGGCAACGCTTACTGAAGAACAG----- 1098
840 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
1099 ----- 1098
900 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 959
1099 -----GTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 1136
960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
1137 GGGCAAACTGTATTCACTCAAAACAAGATCTTTCTTCAACACTGAAGAATAAGTCAGTGAC 1196
1020 CCAGAACACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
1197 CCAGAACACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1256
1080 ACTTGAAGAAGAGTACAGCACAGA 1102
1257 ACTTGAAGAAGAGTACAGCACAGA 1279

RESULT 4
AY399455
LOCUS
DEFINITION
AY399455 5697 bp DNA linear GSS 12-DEC-2003
Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.


```
source      1. .728
/organism="Macaca mulatta"
/mol_type="mrna"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN
Query Match      36.6%; Score 585.2; DB 14; Length 728;
Best Local Similarity 96.9%; Pred. No. 4.6e-122;
Matches 618; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY 295 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGA 354
Db 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAACAGGGAATTATCAGAAGA 60

QY 355 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 414
Db 61 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 120

QY 415 GGTAGCTAGATGGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC- 473
Db 121 GGTAGCTAGATGGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCA 180

QY 474 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGAACAGGAAAATGGA 533
Db 181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAGAAAAGAACAGGAAAATGGA 240

QY 534 GGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACACATAAAGGT 593
Db 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAAAGGT 300

QY 594 GCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATCTCTCACTCACATGGTGGT 653
Db 301 GCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATCTCTCACTCACATGGTGGT 360

QY 654 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 713
Db 361 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 420

QY 714 ATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACA 773
Db 421 ATTGGGAGATCGATGGGCAACATCTGCAGATGGACAGAGACCGCTGGGTCTTTTACA 480

QY 774 AGACATCCTTCTCAAAATGGCAACGTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCT 833
Db 481 AGACATCCTTCTCAAAATGGCAACGTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCT 540

QY 834 TTCAGAAAAAGAGATGCAGTGAACAGATTTCACACAACCT-GGCTTTAAAGATCAAAATG 892
Db 541 TTCAGAAAAAGAGATGTGGTGAACAGGATTCACACAACCTGGGCTTTAAGGATCAAAATG 600

QY 893 AAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAA 930
Db 601 AAATGTTATCAAGTCTTCAAAAACCTGGCCGTTATGTAA 638

RESULT 6
AY399454
LOCUS
DEFINITION Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY399454 genomic survey sequence.
ACCESSION AY399454
VERSION AY399454.1 GI:39755443
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
```

```
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 5676)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5676)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..5676
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5676
/locus_tag="HCM0229"
gene
ORIGIN
Query Match 31.4%; Score 502; DB 29; Length 5676;
Best Local Similarity 75.1%; Pred. No. 5.1e-103;
Matches 634; Conservative 0; Mismatches 101; Indels 109; Gaps 2;

QY 260 GGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGA 319
Db 529 GGGTACATGATGGATNNNNAGCCCATCAGGGCCGGTTGGTAATANNNNNANTTGGGA 588

QY 320 AGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG 379
Db 589 AGTNNNTGATTGGAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG 648

QY 380 ATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGC 439
Db 649 ATGAATCTNNNNNATTCAAGATGGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 708

QY 440 AATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACTCAAAAGAGTTGAATGACTGG 498
Db 709 NNTTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGG 768

QY 499 CTAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAGAGCCTCTTGGACCTGATCTT 558
Db 769 CTAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAGAGCCTCTTGGACCTGATCTT 828

QY 559 GAAGACCTAAACGCCCAAGTACAACACATAAAGGTGCTTCAAGAAGATCTAGAACAAAGAA 618
Db 829 GAAGACCTAAACGCCCAAGTACAACACATAAAGNNNNNNNCAAGAAGATTTAGAACAAAGAA 888

QY 619 CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAT 678
Db 889 CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAANNNNANTGGAGAC 948

QY 679 CACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATC 738
Db 949 CACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATC 1008

QY 739 TGTAGATGGACAGAAAGACCGCTGGTCTTTTACAGACATCCTTCTCAAATGGCAACGT 798
Db 1009 TGTAGGTGGACAGAAAGCCCGCTGGGTCTTTTACAGACATCCTTCTCAAATGGCAACGT 1068

QY 799 CTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGAAAAAAGAGATGCAGTGAAC 858
|||||
```



```
b 1069 CTTACTGAAGAACAG----- 1083
Y 859 AAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACGTG 918
b 1084 ----- 1083
Y 919 GCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGCAAACTGTATTCACTC 978
b 1084 ---GTTTAAAGCGGATNNNGAAAAGAAAAGCAATCCATGGSCAAACTATATTNNNN 1140
Y 979 AAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGACCCAGAAAGACGGAAGCATGG 1038
b 1141 NNACAAGNTNTTCTNTNAACACTGANGAATAAGTCAGTGACCCNNAAGACGGAAGCNGG 1200
Y 1039 CTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAAAACTTGAAAAGAGTACAGCA 1098
b 1201 NNGGATAACTTTGCCCGNNNTTGGGATAATTNNNCCAAAAACNNGAAAAGAGTACAGCA 1260
Y 1099 CAGA 1102
b 1261 CAGA 1264
```

```
ESULT 7
B177816/c
OCUS CBI77816 595 bp mRNA linear EST 31-JAN-2003
EFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; , mRNA sequence.
CCESSION CBI77816
ERSION CBI77816.1 GI:28186206
EYWORDS EST.
OURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 595)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
Location/Qualifiers
1. .595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue type="Purified pancreatic islet"
/lab_host="DH10B"
/clone lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
```

```
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."
ORIGIN
Query Match 31.0%; Score 497; DB 14; Length 595;
Best Local Similarity 99.8%; Pred. No. 4.6e-102;
Matches 508; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGGAGCCCAT 60
Db 509 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGGAGCCCAT 450
QY 61 TCCITTCACAGCATTTGGAGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 449 TCCITTCACAGCATTTGGAGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 390
QY 121 TGAAGTAAACCTGGACCGTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
Db 389 TGAAGTAAACCTGGACCGTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 330
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 329 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 270
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 269 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 210
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAAGATGAAGA 360
Db 209 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAAGATGAAGA 150
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 149 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 90
QY 421 TAGCATGAAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 89 TAGCATGAAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 30
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAA 508
Db 29 GAAAGAGTTGAATGACTGGCTAAACAAAAA 1
RESULT 8
CB850319
LOCUS CB850319 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
DEFINITION MRA-0070
ACCESSION CB850319
VERSION CB850319.1 GI:34380806
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1047)
Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
Swaroop,A.
Annotation and analysis of 10,000 expressed sequence tags from
developing mouse eye and adult retina
Genome Biol. 4 (10), R65 (2003)
22881944
14519200
Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu.
```

FEATURES	Location/Qualifiers	
	source	1..1047
/organism="Mus musculus"		
/mol_type="mRNA"		
/db_xref="taxon:10090"		
/tissue_type="retina"		
/clone_lib="MOUSE ADULT RETINA"		
/note="Vector: pSPORT1"		
ORIGIN		
Query Match 30.1%; Score 481.2; DB 14; Length 1047;		
Best Local Similarity 83.1%; Pred. No. 2e-98;		
Matches 609; Conservative 0; Mismatches 118; Indels 6; Gaps 6;		
2Y	133	GGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTCTGCTGAGGACAC 192
2b	86	GGATAGTTACCAAACTGCTTTANAAGAAGTACTTTTCATGGCTTCTTTCTCCGAGGATAC 145
2Y	193	ATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGACCAGTTTCATAC 252
2b	146	ATTGCCAGCACAAAGGAGAGATTCTCAAATGATGTGGAAGAAGTGAAGAACAGTTTCATGC 205
2Y	253	TCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGCCGGTGTGGTAATATTCTACA 312
2b	206	TCATGAGGGATTTCATGATGGATCTGACATCTCATCAAGGACTTGTGGTAATGTTCTACA 265
2Y	313	ATTGGGAAGTAAGCTGATTGGAAACAGGAAATAATTATCAGAAGATGAAGAACTGAAGTACA 372
2b	266	GTTAGGAAGTCAACTAGTTGGAAAAGGAAATTTATCAGAAGATGAAGAGCTGAAGTGCA 325
2Y	373	AGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAG-CTAGCATGGAAA 431
2b	326	AGAACAAATGAATCTCCTAAATTCAGATGGGAATGTCTCAGGGTANCCTAGCATGGAAA 385
2Y	432	AACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATC-GAAACTGAAAGAGTTGA 490
2b	386	AACAAAGCAAAATTACACAAAGTTCTAATGGATCTCCAGAAATCAGAAATTAAAGAACTAG 445
2Y	491	ATGACTGCTTAACAAAACACAGAAGAAAGAACAGGAAAATGGAGGAGAGCCCTCTTGGAC 550
2b	446	ATGACTGTTTAACAAAAACCTGAAGAGAGAACTAANAANAATGGAGGAGAGCCCTTGGAC 505
2Y	551	CTGATCTTGAAGACCTAAACGCCCAAGTACAAACATAGGTGC-TTCAAGAAGATCTA 609
2b	506	CTGATCTTGAAGATCTAAATGCCAAGTACAAACATAGGTGCTTTCAAGAAGATCTA 565
2Y	610	GAACAAGAACAAAGTCAGGGTCAATCTCTCACTACATGGTGGTGGTAGTTGATGAATCT 669
2b	566	GAACAGGAGCAGGTCAAGGTCAACTCGCTCACTCACTAGNAGTAGTGGGTGATGAATCC 625
2Y	670	AGTGGAGATCAGCAACTGTGTTTGAAGAAACAAC-TTAAGGTATTGGGAGATCGATG 728
2b	626	AGCGGGATCATGCAACAGCTGCTTTGGAANAACAACCTTTAAGGNACTGGGAGATCGATG 685
2Y	729	GGCAAAACATCTG-TAGATGGACAGAAGACCGCTGGGT-CTTTTACAGACATCCTTCTC 786
2b	686	GGCAAAATATCTGCAAAATGGACTGAAAACCGTTGGATTGTTTTTACANGATTTTCTTCTA 745
2Y	787	AAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTACTGTCATGGCTTTCAGAAAAAGAA 846
2b	746	AAATGGCAGCNTTTTACTGAAAAANANNGCCTTTTATNCANGNTTTTAAAAAAGAAAA 805
2Y	847	GATGCAGTGAACA 859
2b	806	TGCATTGANAACA 818
RESULT 9		
BC011062 2334 bp mRNA linear HTC 04-MAR-2003		
LOCUS Mus musculus, Similar to utrophin, clone IMAGE:3979320, mRNA.		
DEFINITION BC011062		
ACCESSION BC011062		
VERSION BC011062.1 GI:15029695		

KEYWORDS	HTC.	
	SOURCE	Mus musculus (house mouse)
ORGANISM		
Mus musculus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE		
1 (bases 1 to 2334)		
AUTHORS Strausberg,R.		
TITLE Direct Submission		
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK		
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov		
Contact: MGC help desk		
Email: cgapbs-r@mail.nih.gov		
Tissue Procurement: Gilbert Smith, Ph.D.		
cDNA Library Preparation: Life Technologies, Inc.		
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center		
Center code: BCM-HGSC		
Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
Contact: amg@bcm.tmc.edu		
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series: IRAX Plate: 23 Row: j Column: 4		
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein		
This clone has the following problem: retained intron.		
FEATURES	Location/Qualifiers	
	source	1..2334
/organism="Mus musculus"		
/mol_type="mRNA"		
/strain="FVB/N"		
/db_xref="taxon:10090"		
/clone="IMAGE:3979320"		
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."		
/clone_lib="NCI CGAP Mam1"		
/lab_host="DH10B"		
/note="Vector: pCMV-SPORT6"		
ORIGIN		
Query Match 20.8%; Score 332.6; DB 11; Length 2334;		
Best Local Similarity 59.0%; Pred. No. 1.2e-64;		
Matches 589; Conservative 0; Mismatches 409; Indels 1; Gaps 1;		
Qy	104	AGTTCAATGATGGAGAGTGAAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTA 163
Db	1042	AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 1101
Qy	164	TTATCGTGGCTTCTTTCTGTGAGGACACATTGCAAGCAAGGAGAGAGATTCTTAATGAT 223
Db	1102	CTGACGTGGCTGTGTCCGGGAGGACACGTTCCAGGAGCAAGATGACATTCTGTATGAT 1161
Qy	224	GTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCC 283
Db	1162	GTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCC 283
Qy	284	CATCAGGGCCGGTGGTAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAA 343
Db	1222	CACCAGAGCAGCGTGGGAGCGTCTGTCAGGCTGGCAACACAGCTGATGACACAGGGACT 1281
Qy	344	TTATCAGAAGATGAAGAACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAGATGG 403
Db	1282	CTGTGAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGTGTAATGCAAGGTGG 1341
Qy	404	GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463

```
b 1342 GAGGCGCTCCGGTGGAGAGCATGGAGAGGCGAGTCCCGGCTGCACGACGCTCTGATGGAG 1401
y 464 CTCCAGAAATCGA-AACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACA 522
b 1402 CTGCAGAAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAAGAGCGCATT 1461
y 523 AGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAA 582
c 1462 CAGAAGATGGAGAGCCTCCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAGCTGCTTCAA 1521
y 583 CAACATAAAGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACT 642
c 1522 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAACT 1581
y 643 CACATGGTGGTGGTGTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAA 702
b 1582 CACATGGTGGTGTGATGGATGAACACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1641
y 703 CAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACGCGCTGG 762
b 1642 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATCCCGCTGGACTGAAGAACGTGG 1701
y 763 GTTCTTTTACAAGACATCCTTCTCAAAATGSCAACGCTTACTGAAGAACAGTGCCCTTTT 822
b 1702 AACAGTTTCAAGAAATCAGTATTCTGTGCGAGGAATTATTGGAAGAGCAGTGTCTGTTG 1761
y 823 AGTGCATGGCTTTACAAAAAAGAAAGATGCAGTGAACAAAGATTACACAACTGGCTTAAA 882
b 1762 GAGGCTTGGCTCACCGAAAGGAAGAGGCTTTGAATAAGTTCAACCCAGCAACTTAAA 1821
y 883 GATCAAAATGAAATGTATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAA 942
b 1822 GACCAGAAGGAACCTAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGA 1881
y 943 AAGAAAAAGCAATCCATGGGCAAACTGTAATCACTCAACAAAGATCTTCTTCAACACTG 1002
b 1882 ATGAAGAGGCAGACTCTGGATCAACTGATGAGATTGGCCAGAGTGTGGGCCAATTACTC 1941
y 1003 AAGAATAAGTCAAGTACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGG 1062
b 1942 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 2001
y 1063 GATAATTTAGTCCAAAACCTTGAAAAGAGTACAGCACAG 1101
b 2002 GATTCTCTGTTTCAGAGACTCGAAGACTCTTCTTAACCAG 2040
```

```
RESULT 10
Y407022
OCUS
EFINITION Homo sapiens UTRN gene, VIRTUAL TRANSCRIPT, partial sequence,
AY407022 genomic survey sequence.
VERSION AY407022.1 GI:39762993
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 9915)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 9915)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
```

```
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..9915
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>9915
/gene="UTRN"
/locus_tag="HCM2752"
ORIGIN
Query Match 20.1%; Score 322.6; DB 29; Length 9915;
Best Local Similarity 62.4%; Pred. No. 3e-62;
Matches 522; Conservative 0; Mismatches 314; Indels 1; Gaps 1;
QY 103 CAGTTCATTGATGGAGAGTGAAGTAAACCTTGACCGTTTATCAAACAGCTTTAGAAGAAGT 162
Db 918 CAGCACTGTCACTGAGGTGACATGGATCTGGACAGCTATCAGATTGCGTTGGAGGAAGT 977
QY 163 ATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCCTAATGA 222
Db 978 GCTGACCTGGTTGCTTCTGCTGAGGACACCTTCCAGGAGCAGGATGATATTTCCTGATGA 1037
QY 223 TGTGAAAGTGGTGAAAGACCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGC 282
Db 1038 TGTTGAAGAAGTCAAAGACCAGTTTGTGCAACCCATGAAGCTTTATGATGGAAGTACTGC 1097
QY 283 CCATCAGGGCCGGTTGGTGAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA 342
Db 1098 ACACCAGAGCAGTGTGGGAGCGTCTCGAGGAGGCAACCACTGATAACACAAGGAAC 1157
QY 343 ATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATG 402
Db 1158 TCTGTCAAGACGAAGAAGATTGAGATTCAGGAACAGATGACCTTGTGAATGCTAGATG 1217
QY 403 GGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGA 462
Db 1218 GGAGGCTCTTAGGGTGGAGAGTATGGACAGACAGTCCCGGCTGCACGATGTCTGATGGA 1277
QY 463 TCTCCAGAA-TCGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAC 521
Db 1278 ACTGCAGAAGAAGCAACTGCAGCAGCTCTCCGCTGGTTAACTCACTCAAGAGGAGCGCAT 1337
QY 522 AAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACA 581
Db 1338 TCAGAAGATGGAAACTTTGCCCTCGATGATGATGATAAAATCTCTACAAAAGCTGCTAGA 1397
QY 582 ACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCAC 641
Db 1398 AGAACATAAAAGTTTGCAAAAGTATCTTGAGGCTGAACAGGTGAAAGTAAATTCACCTAAC 1457
QY 642 TCACATGGTGGTGTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGA 701
Db 1458 TCACATGGTGGTCAATTGTTGATGAAGAACAGTGGTGAGAGTGCTACAGCTATCCTAGAAGA 1517
QY 702 ACAACTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGCTG 761
Db 1518 CCAGTTACAGAAACTTTGGTGAGCGCTGGACAGCAGTATGCCGTTGGACTGAAGAAGCGCTG 1577
QY 762 GGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTCTTACTGAAGAACAGTGCCTTTT 821
Db 1578 GAATAGGTTACAAGAAATCAATATATTGTTGGCAGGAATTATTGGAAGAACAGTGTCTTGT 1637
QY 822 TAGTGCATGGCTTTCAAGAAAAAGAGATGTCAGTGAACAAAGATTTCACACAACTGGCTTTAA 881
Db 1638 GAAAGCTTGGTTAACCCGAAAAAAGAGAGGCTTTTAAATAAAGTCCAGACACAACTTCAA 1697
QY 882 AGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCT 938
```


1698 AGACCAAAGGAACAACTAAGTGTCAAGTGTTCGACGCTGGCTGTAAGTGATGGGTTGT 1754

AY407024 9691 bp DNA linear GSS 15-DEC-2003

Mus musculus UTRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY407024
AY407024.1 GI:39762995

GSS.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 9691)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

14671302
2 (bases 1 to 9691)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..9691
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>9691
/gene="UTRN"
/locus_tag="HCM2752"

ORIGIN

Query Match 18.9%; Score 301.8; DB 29; Length 9691;
Best Local Similarity 61.3%; Pred. No. 1.5e-57;
Matches 503; Conservative 0; Mismatches 317; Indels 1; Gaps 1;

QY 104 AGTTCATTGATGAGAGTGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGTA 163
DB 919 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACAGATAGCGCTAGAGGAGTG 978

QY 164 TTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGAT 223
DB 979 CTGACGTGGCTGCTGCTCCGGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1038

QY 224 GTGGAAGTGGTGAAGACCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCC 283
DB 1039 GTCGAAGAAGTCAAGAGAGCAGTTTGTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1098

QY 284 CATCAGGCGCGGTGGTGGTAATATTCTCAATTTGGGAAGTAAGCTGATTGGAACAGGAAAA 343
DB 1099 CACCAGAGCAGCGTGGGAGCGTCTCGAGGCTGGCAACCAGCTGATGACACAAGGGACT 1158

QY 344 TTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGG 403
DB 1159 CTGTCAAGAGGAGGAGGAGTGTGATGCCAGGAACAGATGACCTGTGATGCAAGGTGG 1218

QY 404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGAT 463
DB 1219 GAGGCGCTCCGGTGGAGAGCATGGAGGCGAGTCCCGGCTGCACGACGCTCTGATGGAG 1278

QY 464 CTCGAGATCGA-AACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACA 522
DB 1279 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAGAGCGCAT 1338

QY 523 AGGAAAATGAGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAA 582
DB 1339 CAGAAGATGAGAGAGCCTCCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAGCTGCTTCAA 1398

QY 583 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAAACAAAGTCAAGTCAATTTCTCTCACT 642
DB 1399 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAACT 1458

QY 643 CACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGAA 702
DB 1459 CACATGGTGGTGTATGTGATGAAGAAACAGTGGGAGAGTGCACAGCTCTTCTTGGAAAGAT 1518

QY 703 CAACCTTAAGGTATGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGG 762
DB 1519 CAGTTACAGAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 1578

QY 763 GTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTT 822
DB 1579 AACAGTTTGCAGAAATCAGTATTCTGTGGCAGGAATATTGGAAGAGCAGTGTCTGTG 1638

QY 823 AGTGCATGGCTTTCAGAAAAAGAGATGAGTGAACAAAGATTCACAACTGGCTTTAAA 882
DB 1639 GAGGCTTGGCTCACCGAAAAGGAAGAGGCTTTGAATAAAGTCAAACCCAGCAACTTAAA 1698

QY 883 GATCAAAATGAATGTTATCAAGTCTTCAAAAACTGGCCGT 923
DB 1699 GACCAGAAGGAACCTAAGTGTGAGTGTCCGGCGTCTGGCTGT 1739

RESULT 12
AY407023 9096 bp DNA linear GSS 15-DEC-2003

LOCUS
Pan troglodytes UTRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY407023
AY407023.1 GI:39762994

GSS.
Pan troglodytes (chimpanzee)

Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 9096)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

14671302
2 (bases 1 to 9096)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..9096
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>9096
/gene="UTRN"
/locus_tag="HCM2752"

ORIGIN

gene

Query Match	18.7%;	Score 299.8;	DB 29;	Length 9096;	
Best Local Similarity	58.4%;	Pred. No. 4.4e-57;			
Matches 489;	Conservative 0;	Mismatches 347;	Indels 1;	Gaps 1;	
/	103	CAGTTCAATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCCTTTAGAAGAAGT	162		
/	918	CAGCACTGTCACCTGAGGTTGACATGGATCTAGACAGCTATCAGATTGCGTTGGAGGAAGT	977		
/	163	ATTATCGTGGCTTCTTCTGCTGAGGACACATTTGCAAGCACAAGGAGAGATTCTAATGA	222		
/	978	GCTGACCTGTTGCTTTCTGCCGAGGACACTTCCAGGAGCAGGATGATATTCTGATGA	1037		
/	223	TGTGGAAGTGGTGAAGACCAGTTTTCATACTCATGAGGGGTACATGATGGATTGACAGC	282		
/	1038	TGTTGAAGAACTCAAGACCAGTTTTCGAACCCATGAANNNNNNNNNNNNNNNNNNNN	1097		
/	283	CCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA	342		
/	1098	NNNNCAGAGCAGTGTGGGCGAGCTCTGCAGGCAGGCCAACCACTGATAANNNAAGGAAC	1157		
/	343	ATTATCAGAAATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATG	402		
/	1158	TCTGTCAGANGAAGAGGAATTTGAGATTCAGGAACAGATNNNNCTGCTGAATGCTAGATG	1217		
/	403	GGAAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGA	462		
/	1218	GGAGGCTCTTAGGGTGGAGAGTATGGACANNAGTCCCGGTGCACGATGTGCTGATGGA	1277		
/	463	TCTCCAGAA-TCGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAAC	521		
/	1278	ACTGCAGAAAGCAACTGCAGCAGCTCTCCGCTGTTAAACACTCACAGAGGAGCGCAT	1337		
/	522	AAGGAAAATCGAGGAAGAGCCCTCTTGGACCTGATCTTTGAACACCTAAACGCCAAGTACA	581		
/	1338	TCAGAAAGATGGAACACTTGCCTGATGATGATGAATCTCTACAAAGCTGCTAGA	1397		
/	582	ACAACATAAGTGCTTCAAGAAGATCTAGAACAAAGAACAACTCAGGGTCAATTTCTCTCAC	641		
/	1398	AGAACATAAAGTTTGCAAGTATCTTGAGGCTGAACAGGTGAAGTAAATTCACATAAC	1457		
/	642	TCACATGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGA	701		
b	1458	TCACATGGTGGTTCATTTGATGAAAACAGTGGTGAGAGTGTCTACAGCTATCCTAGAAGA	1517		
/	702	ACAACTTAAGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTG	761		
b	1518	CCAGTTACAGNNNNNNNNNNNGCTGGACAGCAGTATGCCGTTGGACTGAAGACGCTG	1577		
/	762	GGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTT	821		
/	1578	GAATAGTTTACAAGAAATCAATATATTGTGGCAGGAATTTATTGGAAGAACAGTGTGTT	1637		
/	822	TAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAA	881		
b	1638	GAAAGCTTGGTTAAACCGAAAAAGAGAGGCTTTAAATAAAGTCCAGACAAGCAACTTCAA	1697		
/	882	AGATCAAAATGAATGTTATCAAGTCTTCAAAAACTGGCGCTTTTAAAGCGGATCT	938		
b	1698	AGACCAAAAGGAACCTAAGTGTCAAGTGTTCGACGCTCTGGCTGTAAAGTATGGGGTTGT	1754		

ESULT 13
F109978
OCUS
EFINITION
Shultzomica03229 Rat lung airway and parenchyma cDNA libraries
Rattus norvegicus cDNA clone Contig2846 5', mRNA sequence.
CCESSION
F109978
ERSON
EYWORDS
SOURCE
Rattus norvegicus (Norway rat)
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1122)
Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
Plopper,C.G. and Buckpitt,A.R.
Gene Expression Analysis in Response to Lung Toxicants: I.
Sequencing and Microarray Development
Unpublished (2003)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 1122.
Location/Qualifiers
1..1122
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig2846"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/note="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

ORIGIN

Query Match	16.5%;	Score 264.8;	DB 14;	Length 1122;	
Best Local Similarity	61.7%;	Pred. No. 2.6e-49;			
Matches 439;	Conservative 0;	Mismatches 272;	Indels 1;	Gaps 1;	
QY	103	CAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCCTTTAGAAGAAGT	162		
Db	285	CAGCACAGTCAACGAGGTGACATGGATTGGACAGCTACCAGATAGCAGTACAGGAAAGT	344		
QY	163	ATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGA	222		
Db	345	ACTGACATGGCTGCTGCTGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGA	404		
QY	223	TGTGGAAGTGGTGAAGACCAAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGC	282		
Db	405	TGTCGAGGACGTCAAAGAGCAGTTTGTACCCATGAAACATTTATGATGGAGCTGAGTGC	464		
QY	283	CCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA	342		
Db	465	GCACCAGAGCAGCGTGGGAGTGTCTCTGCAGGCTGGCAACCACTGATGACACAAGGCAC	524		
QY	343	ATTATCAGAAGATGAAGAACTGAAGTACACAGCAGATGAATCTCTTAAATTCAGATG	402		
Db	525	TCTGTCGGATGAGGAGGAGTTTGAGATTCAGGAGCAGATGACCTTGTGAACGCCAGATG	584		
QY	403	GGAAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGA	462		
Db	585	GGAGGCGCTCCGGGTAGAGAGCATGGAGAGACAGTCCCGTTGCATGATGCCCTGATGGA	644		
QY	463	TTCCAGAAATCGA-AACTGAAAGAGTTGATGATGCTGCTAACAACAAACAGAGAAGAAC	521		
Db	645	ACTGCAAGAAACACAGCTGCAGCAGCTCTCAGGCTGGCTGACCCCTCACAGAGGCGCAT	704		
QY	522	AAGGAAAATGGAGGAGAGCCTCTTGGACCTCTTGAAGACCTTAAACGCCCAAGTACA	581		
Db	705	TCAGAAGATGGAAGCCCTCCAGTGGGTGACGACCTACCCCTCCCTGCAAAACCTGCTTGA	764		
QY	582	ACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGCAAGTCAGGTCATTTCTCTCAC	641		

Db 765 AGACATATAAAAGTTTGCAAAGTGACCTCGAAGCTGAGCAGGTGAAGGTGAATTCCTTAAC 824
Y 642 TCACATGGTGGTGGTAGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGA 701
Db 825 TCATATGGTGGTGAATTTGGATGAAACACAGTGGGAGAGCGCCACAGCTGTTTGAAGA 884
Y 702 ACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTG 761
Db 885 TCAGTTACAGAAACTGGGTGACCGCTGGACAGCTCTATGCCGCTGGACTGAAGAACGTTG 944
Y 762 GGTTCCTTTACAAGACATCCTTCTCAAAATGGCAACGCTTTACTGAAGAACAG 813
Db 945 GAACAGGTTGCAAGAAATCAATATTTTGTGGCAGGAATTATTGGAAGAACAG 996

RESULT 14
BI250598 772 bp mRNA linear EST 17-JUL-2001
LOCUS 6029936559f1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149431 5',
DEFINITION mRNA sequence.
ACCESSION BI250598
VERSION BI250598.1 GI:14799101
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 772)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11369 row: e column: 16
High quality sequence start: 3
High quality sequence stop: 771.

FEATURES
source
1..772
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5149431"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: Not1; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN
Query Match 16.0%; Score 255.4; DB 12; Length 772;
Best Local Similarity 61.4%; Pred. No. 3.3e-47;
Matches 460; Conservative 0; Mismatches 286; Indels 3; Gaps 3;

QY 104 AGTTCATTGATGGACAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAGTA 163
Db 26 AGCACCGTCACCTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 85
QY 164 TTATCGTGGCTTCCTTCTGCTGAGGACACATTGCAAGCAAGCAGAGAGATTTCTAATGAT 223
Db 86 CTGACGTGGCTGCTGCTCCGGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 145

QY 224 GTGGAAGTGGTGAAAGACCAGCTTTCATCTCATGAGGGGTACATGATGGATTGCACAGCC 283
Db 146 GTCGAAGAAGTCAAAGAGCAGTTGC-TACCCATGAAGACTTTTATGATGGAGCTACAGCA 204
QY 284 CATCAGGCGCGGTTGGTAATATCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAA 343
Db 205 CACCAGAGCAGCGTGGGAGCGTCTGCAGGCTGGCAACACAGCTGATGACACAAGGGACT 264
QY 344 TTATCAGAAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGG 403
Db 265 CTGTCAAGAGGAGGAGCAGTTGAGATCCAGGAACAGATGACCTTGTGTAATGCAAGGTGG 324
QY 404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
Db 325 GAGGCGCTCCGGGTGGAGAGCATGGAGAGCGAGTCCCGGCTGCACGACGCTCTGATGGAG 384
QY 464 CTCCAGAAATCGA-AACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACA 522
Db 385 CTGCAGAAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAGAGCGCATT 444
QY 523 AGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAA 582
Db 445 CAGAAGATGGAGAGCCTCCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAGCTGCTTCAA 504
QY 583 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTTCTCTCACT 642
Db 505 GAACATAAAATTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAACT 564
QY 643 CACATGGTGGTGGTAGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAA 702
Db 565 CACATGGTGGTGGTAGTGGATGAAAAACAGTGGGGAGAGTGCACAGCTCTTCTGGAAGAT 624
QY 703 CAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGG 762
Db 625 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 684
QY 763 GTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTCTACTGAAGAACAGTGCCTTTT 822
Db 685 AACAGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTC-TGTA 743
QY 823 AGTGCATGGCTTTCAGAAAAAGAGATGC 851
Db 744 GGAGGCTGGATCACCGAAAGGAAGAGGC 772

RESULT 15
AK081426 3753 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
DEFINITION library, clone:Cl30016K19 product:dystrophin related protein 2,
full insert sequence.

ACCESSION AK081426
VERSION AK081426.1 GI:26349154
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model
un on: April 4, 2004, 16:58:13 ; Search time 77.6428 Seconds
(without alignments)
11443.127 Million cell updates/sec

itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgctctacacacagg.....taccagtttttagtgacct 1601

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

eached: 682709 seqs, 277475446 residues

otal number of hits satisfying chosen parameters: 1365418

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Match	Length	DB ID	Description
1	1092.4	68.2	5952	4	US-09-687-875A-1
2	1091	68.1	13977	4	US-09-484-970B-60
C 3	854.2	53.4	19307	3	US-08-836-022A-10.
C 4	854.2	53.4	19307	3	US-09-427-048A-10
5	332.6	20.8	6045	4	US-09-091-501B-7
6	332.6	20.8	10320	4	US-09-091-501B-9
7	234.2	14.6	3915	4	US-09-976-594-93
8	79.4	5.0	200	4	US-09-091-501B-5
9	78.6	4.9	200	4	US-09-091-501B-4
10	78.6	4.9	200	4	US-09-091-501B-6
C 11	74.6	4.7	7218	1	US-08-232-463-14
12	42.8	2.7	1179	4	US-09-107-532A-1186
13	42.8	2.7	1690	4	US-09-620-312D-69
14	42.8	2.7	2574	4	US-09-668-313A-10
15	42.8	2.7	7812	3	US-09-368-590-1
C 16	42.4	2.6	2704	3	US-08-857-076-44
C 17	42.4	2.6	3499	3	US-08-857-076-43
C 18	40.6	2.5	505	4	US-09-621-976-15639
19	39.4	2.5	832	4	US-09-621-976-2813
20	39.4	2.5	2223	1	US-08-257-073-4
C 21	39.4	2.5	193303	4	US-09-497-855A-37
C 22	39.4	2.5	193303	4	US-09-497-855A-44
23	38.8	2.4	289	3	US-09-007-005-17
24	38.8	2.4	289	3	US-09-244-796-17
25	38.8	2.4	1821	4	US-08-477-831C-1
26	38.8	2.4	1885	4	US-08-477-831C-9
27	38.8	2.4	1896	4	US-08-477-831C-10

28	38.8	2.4	1961	4	US-08-477-831C-8	Sequence 8, Appli
29	38.8	2.4	2968	4	US-08-477-831C-13	Sequence 13, Appl
30	38.8	2.4	3044	4	US-08-477-831C-12	Sequence 12, Appl
31	38.4	2.4	7672	4	US-09-220-132-24	Sequence 24, Appl
32	38.2	2.4	1848	4	US-09-134-001C-447	Sequence 447, App
C 33	37.4	2.3	2082	3	US-08-985-335-4	Sequence 4, Appli
C 34	37.4	2.3	2082	3	US-09-410-372-4	Sequence 4, Appli
35	37.2	2.3	2160	4	US-09-092-218-1	Sequence 1, Appli
C 36	37.2	2.3	2915	4	US-09-336-115C-5	Sequence 5, Appli
C 37	37.2	2.3	3902	4	US-08-961-527-212	Sequence 212, App
38	36.8	2.3	4929	4	US-09-620-312D-674	Sequence 674, App
39	36	2.2	608	3	US-09-385-982-236	Sequence 236, App
40	36	2.2	2763	1	US-08-248-466B-2	Sequence 2, Appli
41	35.8	2.2	633	4	US-09-134-001C-578	Sequence 578, App
42	35.6	2.2	2447	2	US-09-014-969-14	Sequence 14, Appl
43	35.6	2.2	4868	1	US-08-139-937-12	Sequence 12, Appl
44	35.6	2.2	4868	5	PCT-US93-11310-12	Sequence 12, Appl
45	35.6	2.2	8257	4	US-09-595-684B-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match	68.2%	Score 1092.4;	DB 4;	Length 5952;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1107;	Conservative 0;	Mismatches 6;	Indels 1;	Gaps 1;
QY 1	GAGCTATGCCCTACACACAGGCTGCTTATGTGTCACCACTCTGACCCCTACACGAGCCCATTT 60			
Db 891	GAGCTATGCCCTACACACAGGCTGCTTATGTGTCACCACTCTGACCCCTACACGAGCCCATTT 950			
QY 61	TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGCGCAGTTCATTGATGGAGAG 120			
Db 951	TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGCGCAGTTCATTGATGGAGAG 1010			
QY 121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180			
Db 1011	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1070			
QY 181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240			
Db 1071	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130			
QY 241	CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTGG 300			

Db 1131 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGCGCGGTTGG 1190
QY 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1250
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1310
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1311 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATCAGAAACT 1370
QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539
Db 1371 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1430
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAAATAGGTGCTTCA 599
Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAAATAGGTGCTTCA 1490
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 659
Db 1491 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 1550
QY 660 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 1551 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1611 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1670
QY 780 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTCGCCCTTTTAGTGCAATGGCTTTCAGA 839
Db 1671 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTCGCCCTTTTAGTGCAATGGCTTTCAGA 1730
QY 840 AAAAGAAGATGCAGTGAAACAAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 899
Db 1731 AAAAGAAGATGCAGTGAAACAAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1790
QY 900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1791 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1850
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAAATAAGTCAGTGAC 1910
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1079
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1970
QY 1080 ACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAGA 1113
Db 1971 ACTTGAAAAAGAGTACAGCACAGGAAACTGAAATA 2004

RESULT 2
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program

; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 68.1%; Score 1091; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 1158
QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTTAGAAGAACTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTTAGAAGAACTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCACAAGCAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCACAAGCAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCCGGTTGG 300
Db 1339 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCCGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAAGTGCCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAAGTGCCTTCA 1698
QY 600 AGAAGATCTAGAAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 659
Db 1699 AGAAGATCTAGAAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGSTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGSTATTGGG 1818
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 780 CCTTCTCAAATGSCAACGTCCTTACTGAAGAACAGTCGCCCTTTTAGTGCAATGGCTTTCAGA 839
Db 1879 CCTTCTCAAATGSCAACGTCCTTACTGAAGAACAGTCGCCCTTTTAGTGCAATGGCTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAAACAAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 899

b	1939	AAAAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT	1998
y	900	ATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	959
b	1999	ATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	2058
y	960	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1019
b	2059	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	2118
y	1020	CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAAA	1079
b	2119	CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAAA	2178
y	1080	ACTTGAAAAAGAGTACAGCACAGA	1102
b	2179	ACTTGAAAAAGAGTACAGCACAGA	2201

ESULT 3

S-08-836-022A-10/c
Sequence 10, Application US/08836022A

PAGE# NO: 600133/
GENERAL INFORMATION:

GENERAL INFORMATION:
 APPLICANT: Trustees of the University of Pennsylvania
 APPLICANT: Wilson, James M.
 APPLICANT: Fisher, Krishna J.
 APPLICANT: Chen, Shu-Jen
 APPLICANT: Weitzman, Matthew
 TITLE OF INVENTION: Improved Adenovirus Virus and
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/836.022A

FILING DATE:

CLASSIFICATION: 4355

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/331,381

FILING DATE: 28-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31.215

REFERENCE/DOCKET NUMBER: GNVEN.008PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-922

TELEPHONE: 215-540-5818
TELEFAX: 215-540-5818

TELEFAX: 212-340-2018
INFORMATION FOR SEQ ID NO: 10:

INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:

SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs

PENICILLIN: 19307 Dase
 TYPE: nucleic acid

TYPE: NUCLEAR ACID
STRANDEDNESS: DOUBLE

STRADEDNESS: 000
TOBOLGY: 1234567890

TOPOLOGY: UNKNOWN
MOLECULE TYPE: cDNA

5-08-83-0329-10
FBI DETROIT

Query Match 53.4%; Score 854.2; DB 3; Length 19307;

Best Local Similarity 86.5%: Pred. No. 1.2e-252:

Sequence similarity	Insertions	Deletions	Indels	Gaps
Matches 954: Conservative	0: Mismatches	148:		

1 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGACCCCTACCGAGCCCAT 60

13568	Db	GAGTTATGCCTTCACACAGGCTGCTTATGTTGGCCACTCTGTATCCACACAGAGCCCTTA	13500
61	Qy	TCCTTCACAGCATTTTGGAACTCCTGAAGACAAGTCAATTTGGCAGTTTCATTTGATGGAGAG	120
13508	Db	TCCTTCACAGCATTTTGGAACTCCAGAGACAAGTCACTTGACAGTTTCATTTGATGGAGAC	13449
121	Qy	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGATATTATCGTGGCTTCCTTC	180
13448	Db	GGAAGTAAATCTGGATAGTTACCAACTGCTTTTAGAAGAGTACTTTTCATGGCTTCCTTC	13389
181	Qy	TGCTGAGGACACATTCGAAGACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA	240
13388	Db	TGCCGAGGATACATTCGGAGACAAGGAGAGATTTTCAAAATGATGTTGAAGAACTGAAAGA	13329
241	Qy	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGTTGG	300
13328	Db	ACAGTTTCATGCTCATGAGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGTGG	13269
301	Qy	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA	360
13268	Db	TAATGTTCTACAGTTAGGAAGTCAACTAGTTTGGAAAAAGGAAAAATTTATCAGAAGATGAAGA	13209
361	Qy	AACGTAAGTACAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC	420
13208	Db	AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAGATGGGAATGTCTCAGGGTAGC	13149
421	Qy	TAGCATGGAATAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATC-GAAACT	479
13148	Db	TAGCATGGAATAACAAAGCAATTTACAAAGTTCTAATGGATCTCCAGAATCAGAAAT	13089
480	Qy	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACAAAGAAATGGAGGAAGA	539
13088	Db	AAAAGAACTAGATGACTGGTTAACAAAACTGAAGAGAGAACTAAGAAAAATGGAGGAAGA	13029
540	Qy	GCCTTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACATTAAGGTGCTTCA	599
13028	Db	GCCCTTTGGACCTGATCTTGAAGATCTAAATGCCCAAGTACAAACATTAAGGTGCTTCA	12969
600	Qy	AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATCTCTCACTCACATGGTGGTGTAGT	659
12968	Db	AGAAGATCTAGAACAGGACGGTCAGGGTCAACTCGCTCACTCACATGGTGTAGTGTGT	12909
660	Qy	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	719
12908	Db	TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGGAAACAACTTAAGGTACTGGG	12849
720	Qy	AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGGTGGTTCCTTTTACAAGACAT	779
12848	Db	AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGGTGGATGTTTTACAAGATAT	12789
780	Qy	CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA	839
12788	Db	TCTTCTAAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA	12729
840	Qy	AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT	899
12728	Db	AAAAGAGATGCAATGAAGAACATTCAGACAAGTGGCTTTTAAAGATCAAAATGAAATGAT	12669
900	Qy	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	959
12668	Db	GTCAAGTCTTCAAAAAATATCTACTTTTAAAAATAGATCTAGAAAAAGAAAAAGCCACCAT	12609
960	Qy	GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC	1019
12608	Db	GGAAAACTAAGTTCACTCAATCAAGATCTACTTTTCGGCACTGAAAAATAAGTCAGTGAC	12549
1020	Qy	CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA	1079
12548	Db	TCAAAAGATGGAATCTGGATGGAAAACTTTTGCAACACGTTGGGACAATTTAACCCAAA	12489
1080	Qy	ACTTGAAGAGGTACAGCACAGA	1102
12488	Db	ACTTGAAGAGGTTCAGCACAAA	12466

RESULT 4
JS-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Wilson, James M.
; Fisher, Krishna J.
; Chen, Shu-Jen
; Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Query Match 53.4%; Score 854.2; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 1.2e-252;
Matches 954; Conservative 0; Mismatches 148; Indels 1; Gaps 1;
2Y 1 GAGCTATGCTACACACAGGCTGCTTATGTACACCCTCTGACCCCTACACGGAGCCCCATT 60
13568 GAGTTATGCTTCCACACAGGCTGCTTATGTTGCCACCTCTGATTCACACAGAGCCCCCTA 13509
2Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATGTATGGAGAG 120
13508 TCCTTCACAGCATTTGGAAGCTCCACAGACAAAGTCACTTGACAGTTCAATGTATGGAGAC 13449
2Y 121 TGAAGTAACTGGACCGCTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 180
13448 GGAAGTAACTGGATAGTTTACCAAACTGCTTTTAGAGAAAGTACTTTTCATGGCTTCTTTC 13389
2Y 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
13388 TGCCGAGGATACATTCGAGCACAAGGAGAGATTTCAATGATGTTGAAGAAGTGAAGA 13329
2Y 241 CCAGTTTCACTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGTGG 300

Db 13328 ACAGTTTTCATGCTCATGAGGGATTTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 13269
QY 301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 360
13268 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAAGGAAAAATTTATCAGAAGATGAAGA 13209
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGTTAGC 420
Db 13208 AGCTGAAGTCAAGAACAAATGAATCTCCTAAATTCAGATGGGAATGTCTCAGGTTAGC 13149
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
13148 TAGCATGGAAAAACAAAGCAATTTACACAAAGTTCTTAATGGATCTCCAGAAATCAGAAAT 13089
QY 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539
Db 13088 AAAAGAACTAGATGACTGGTTAAACAAAACTGAAGAGAGAACTAAAGAAAAATGGAGGAAGA 13029
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGSTGCTTCA 599
Db 13028 GCCCTTTGGACCTGATCTTGAAGATCTAAATGCCCAAGTACAAACATAAGSTGCTTCA 12969
QY 600 AGAAGATCTAGAACAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 659
12968 AGAAGATCTAGAACAAGGAGCAGGTCAAGGTCAACTCGTCACTCACATGGTGTAGTGGT 12909
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719
Db 12908 TGATGAATCCAGGGTGTATCATGCAACAGCTGCTTTTGAAGAACAACTTAAGGTACTGGG 12849
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACAGCCGCTGGGTTCTTTTACAGACAT 779
12848 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCCGCTGGATTGTTTTACAGATAT 12789
QY 780 CCTTCTCAAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTTCA 839
12788 TCTTCTAAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTAGTACATGGCTTTTCA 12729
QY 840 AAAAGAAAGATGCAGTGAAACAAGATTCAACAACTGGCTTTAAAGATCAAAATGAATGTT 899
Db 12728 AAAAGAAAGATGCAATGAAGAACATTCAGACAAAGTGGCTTTAAAGATCAAAATGAATGAT 12669
QY 900 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
12668 GTCAAGTCTTCAAAAAATATCTACTTTAAAAATAGATCTAGAAAAAGAAAAAGCCCAACCAT 12609
QY 960 GGGCAAACTGTATTCTCACTCAAAACAAGATCTTTCTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 12608 GGAATAAACTAAGTTCACTCAATCAAGATCTACTTTCCGCACTGAAAAATAAGTCAGTGAC 12549
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGCTTGGGATAATTTAGTCCAAAA 1079
12548 TCAAAAGATGGAAATCTGGATGGAAAACTTTTGCAACAGTTGGGACAATTTTACCCAAAA 12489
QY 1080 ACTTGAAGAGAGTACAGCACAGA 1102
Db 12488 ACTTGAAGAGAGTTCAGCACAAA 12466

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8

PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
S-09-091-501B-7
Query Match 20.8%; Score 332.6; DB 4; Length 6045;
Best Local Similarity 59.0%; Pred. No. 8.2e-92;
Matches 589; Conservative 0; Mismatches 409; Indels 1; Gaps 1;
Y 104 AGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGTA 163
b 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973
Y 164 TTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTTAATGAT 223
b 974 CTGACGTGGCTGCTGCTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033
Y 224 GTGGAAGTGGTGAAGAACACAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCC 283
b 1034 GTCGAAGAAGTCAAGAGAGCAGTTTGTACCCATGAACACTTTATGATGGAGCTGACAGCA 1093
Y 284 CATCAGGCGCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAA 343
b 1094 CACCAGAGCAGCGTGGGGAGCGTCTCGAGGCTGGCAACCAGCTGATGACACAAGGACT 1153
Y 344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGG 403
b 1154 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGCTGAATGCAAGGTG 1213
Y 404 GAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTCTTAATGGAT 463
b 1214 GAGGCGCTCCGGGTGAGAGCATGGAGAGGAGCTCCGGCTGCACGACGCTCTGTATGGAG 1273
Y 464 CTCCAGAATCGA-AACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACA 522
b 1274 CTGCAGAAGAAACAGTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAGAGCGCATT 1333
Y 523 AGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAA 582
b 1334 CAGAAGATGGAGAGCCCTCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAGCTGCTTCAA 1393
Y 583 CAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGAAACAAGTCAGGTCATTTCTCTCACT 642
b 1394 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAGGTAAATTCCTTAACT 1453
Y 643 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCCAATCTGTAGTGCAGAACCCGCTGG 702
b 1454 CACAIGGTGGTGATTGTGGATGAAAACAGTGGGAGAGTGCCACAGCTCTTCTTGGAGAT 1513
Y 703 CAACCTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGTGCAGAACCCGCTGG 762
b 1514 CAGTTACAGAAACTGGGTAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 1573
Y 763 GTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTT 822
b 1574 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATATTGGAAGAGCAGTGTCTGTG 1633

QY 823 AGTGCATGGCTTTTCAGAAAAAAGAAAGATGCAGTGAACAAGATTTCACACAACACTGGCTTTTAAA 882
Db 1634 GAGGCTTGGCTCACCGAAAGAAAGAGAGGCTTTGAATAAAGTTTCAAAACCCAGCAACTTTAAA 1693
QY 883 GATCAAAATGAAATGTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAA 942
Db 1694 GACCAGAAGGAACATAAGTGTCACTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1753
QY 943 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTG 1002
Db 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGCCCAATTACTC 1813
QY 1003 AAGAATAAGTCACTGACCCAGAAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGG 1062
Db 1814 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 1873
QY 1063 GATAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAG 1101
Db 1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTTAACCAG 1912

RESULT 6
US-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 20.8%; Score 332.6; DB 4; Length 10320;
Best Local Similarity 59.0%; Pred. No. 1.2e-91;
Matches 589; Conservative 0; Mismatches 409; Indels 1; Gaps 1;
QY 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTA 163
Db 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGTAGAGGAAGTG 973
QY 164 TTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTTAATGAT 223
Db 974 CTGACGTGGCTGCTGTCCCGGAGGACACCGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033
QY 224 GTGGAAGTGGTGAAGAACCGAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCC 283

Db 1034 GTCGAAGAAAGTCAAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1093
QY 284 CATCAGGCGCGGTTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAA 343
Db 1094 CACCAGAGCAGCGTGGGAGCGCTCTGCAGGCTGGCAACAGCTGATGACACAAGGGACT 1153
QY 344 TTATCAGAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
Db 1154 CTGTACAGAGGAGGAGGAGTTTGATATCCAGGAACAGATGACCTTGCTGAATGCAAGGTGG 1213
QY 404 GAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
Db 1214 GAGGCGCTCCGGTGGAGAGCATGGAGAGGCAGTCCCGGCTGCACGACGCTCTGATGGAG 1273
QY 464 CTCAGAAATCGA-AACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACA 522
Db 1274 CTGCAGAAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAGAGCGCAT 1333
QY 523 AGGAAATGGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAAGTACAA 582
Db 1334 CAGAAGATGGAGAGCCCTCCGCTGGTGATGACCTGCCCTCCCTGCAGAAAGCTGCTTCAA 1393
QY 583 CAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTACT 642
Db 1394 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAACT 1453
QY 643 CACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAA 702
Db 1454 CACATGGTGGTAGTTGGATGAAACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1513
QY 703 CAACTTAAGGTATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGG 762
Db 1514 CAGTTACAGAAACTGGGTGAGCGCTGCACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 1573
QY 763 GTTCTTTTACAGACATCCTTCTCRAATGGCAACGTCCTACTGAAGAACAGTGCCTTTT 822
Db 1574 AACAGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTCTGTTG 1633
QY 823 AGTGCATGGCTTTCAGAAAAAGAAAGATGCAGTGAAACAAGATTACACAACTGGCTTTAAA 882
Db 1634 GAGGCTTGCTACCCGAAAGAAAGAGAGGCTTTGAATAAAGTTCAAAACAGCAACTTTAAA 1693
QY 883 GATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAA 942
Db 1694 GACCAGAAGGAAGTAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1753
QY 943 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAAGATCTTCTTTCAACACTG 1002
Db 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTACTC 1813
QY 1003 AAGAATAAGTCAGTGACCCAGAACGGAAGCATGGCTGGATAAATTTGCCGCTGTTGG 1062
Db 1814 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGAATCTGAGGAGCTAACACAGAGATGG 1873
QY 1063 GATAATTTAGTCCAAAAAAGTTGAAAAGAGTACAGCACAG 1101
Db 1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTTAACCAG 1912

RESULT 7
US-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143

; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match 14.6%; Score 234.2; DB 4; Length 3915;
Best Local Similarity 58.6%; Pred. No. 1.6e-61;
Matches 323; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1131 GGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCA 1190
Db 1022 GGGCATGGAGGAACTAAGCACTACTCTGAGCCAAAGCTGAGGAGTCCGAGCCACTTGGGA 1081
QY 1191 GCCCCTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACT 1250
Db 1082 GCCCATTTGGGATCTCTTCAATTGATTCACTCCAGAGCACATCCAGGCTATTAAAGCTGTT 1141
QY 1251 TCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTGGCA 1310
Db 1142 CAAAGAAGAAATTTCCCCCATGAAGATGGAGTAAAGTTGTGTAATGATCTGGCCCCACCA 1201
QY 1311 GCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAA 1370
Db 1202 ACTTGCCATTTCTGATGTGCACCTTGTCAATGGAGAAATTTCCAGGCCCTTGAACAGATCAA 1261
QY 1371 CACCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGC 1430
Db 1262 CGTCCGATGGAACAACACTACAGGCGTCACTTGTGATGAGAGGCTTAAGCAGCTCCAGGATGC 1321
QY 1431 CCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCTTTCCAGCTGTCTCCAGGGTCCCTG 1490
Db 1322 CCACCGGACCTTTGGCCCTGGGTCAAGCACTTTTCTCTCTCTCTGTCCAGGTTCCTG 1381
QY 1491 GGAGAGAGCCATCTGCCAAACAAAGTGCCTTACTATATCAACCAGAGACTCAAAACAC 1550
Db 1382 GGAAGAGCAATTCACCCCAATAAAGTTCCCTACTACATCAACCACAGGCTCAGACCAC 1441
QY 1551 TTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
Db 1442 ATGCTGGACCATCCCAAGATGACAGAGTTATACCAAAACCCCTAGCTGATCT 1492

RESULT 8
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 69
LENGTH: 1690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(1522)
S-09-620-312D-69

Query Match 2.7%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.012;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
Y 1116 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 1175
b 55 CAAGGAGTTGCACCAAGTGGCGCACGACCTGGACGACGAGCTGGCATGGTTTCAGGAGCG 114
Y 1176 CAAGGGA---TCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCT 1232
b 115 GCTGCCACTGGCCATGCAGACAGAGCGGACACGGTTTGCAGCGGTCCAGCAGACAT 174
Y 1233 CGAGAAAGTCAAGGCACCTTCAGGAGGAATTCGGCCTCTGAAAGAGAACGTGAGCCAGT 1292
b 175 CAAAAGAACCAAGGCGCTCGCGGGAGATCCAGGCGCATGGGCCGCGCTGGAGAGGT 234
Y 1293 CAATGACCTTGCTGCCAGCTTACCATTGGGCATTGAGCTTACCGTATACCGTATACCTCAG 1352
b 235 GCTGGAGCGCGGGCGCGCTGGCGTGGCTGCGCAGCCCGGAGGCAGAGGCAGTGGCGCG 294
Y 1353 CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGT 1412
b 295 GGGCCTGGAGCAGCTGCAGAGCGCCTGGCGCGGACTGCGGAGGCTGCCGAGCGACGGCA 354
Y 1413 CAGGCAGCTGCATGAAGCCACAGGG 1438
b 355 GCAGGTGCTGGACGCGCGCTTCCAGG 380

ESULT 14
S-09-668-313A-10
Sequence 10, Application US/09668313A
Patent No. 6503756
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Susan M. Freier
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
FILE REFERENCE: RTS-0127
CURRENT APPLICATION NUMBER: US/09/668,313A
CURRENT FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 247
SEQ ID NO 10
LENGTH: 2574
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (218)...(1891)

US-09-668-313A-10
Query Match 2.7%; Score 42.8; DB 4; Length 2574;
Best Local Similarity 61.8%; Pred. No. 0.016;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1489 TGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCTACTATATCAACACGAGACTCAAACA 1548
Db 1733 TGGGAGGAAGCTTACACAGCAGATGGAATCAAGTACTTTCATCAACACGAGTGCACAGACC 1792
QY 1549 ACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGCTGA 1598
Db 1793 ACGTCTGGATCCACCCGCTGATGAGCGCCCTGAACCTGTCTGTGCAGA 1842
RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1
Query Match 2.7%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.036;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1116 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAGCTGAGGTGAT 1175
Db 3639 CAAGGAGTTGCACCAAGTGGCGCACGACCTGGACGACGAGCTGGCATGGGTTCAGGAGCG 3698
QY 1176 CAAGGGA---TCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCT 1232
Db 3699 GCTGCCACTGGCCATGCAGACAGAGCGGACACGGTTTGCAGGCGGTCCAGCAGCACAT 3758
QY 1233 CGAGAAAGTCAAGGCACCTTCAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGT 1292
Db 3759 CAAAAGAACCAAGGCGCTCGCGGGAGATCCAGGCGCATGGGCCGCGCTGGAGGAGGT 3818
QY 1293 CAATGACCTTGCTGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAG 1352
Db 3819 GCTGGAGCGCGGGCGCGCTGGCGTGGCTGCGCAGCCCGGAGGCGCAGAGGCAGTGGCGCG 3878
QY 1353 CACTCTGGAAGACCTGAACACCAGATGGAAGTCTTGCAGGTGGCCGTCGAGGACCGAGT 1412
Db 3879 GGGCCTGGAGCAGCTGCAGAGCGCCTGGGCCGAGCTCGGGAGGCTGCGGAGCGGACGGCA 3938
QY 1413 CAGGCAGCTGCATGAAGCCACAGGG 1438
Db 3939 GCAGGTGCTGGACGCGCGCTTCCAGG 3964

Search completed: April 5, 2004, 08:25:02
Job time : 79.6428 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
nucleic - nucleic search, using sw model
on on: April 4, 2004, 11:55:32 ; Search time 413.121 Seconds
(without alignments)
16463.377 Million cell updates/sec

itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgcctacacacagg.....taccagttcttagctgacct 1601
oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
earched: 3373863 seqs, 2124099041 residues
otal number of hits satisfying chosen parameters: 6747726

inimum DB seq length: 0
aximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					%		Query		Match		Length		DB		ID		Description	
result	No.	Score	Match	Length	DB	ID												
1	1601	100.0	3446	6	AAD37242	Aad37242 Human dys												
2	1590	99.3	4414	6	AAD37260	Aad37260 Adeno-ass												
3	1420.8	88.7	5417	6	ABK81997	Abk81997 DNA encod												
4	1193	74.5	3510	6	AAD37240	Aad37240 Human dys												
5	1193	74.5	4476	6	AAD37259	Aad37259 Adeno-ass												
6	1102	68.8	4182	6	AAD37230	Aad37230 Human dys												
7	1102	68.8	5149	6	AAD37255	Aad37255 Adeno-ass												
8	1092.4	68.2	5952	5	AAD06794	Aad06794 Human dys												
9	1091	68.1	8689	6	ABK82000	Abk82000 DNA encod												
10	1091	68.1	11241	6	ABK82005	Abk82005 cDNA encod												
11	1091	68.1	11443	6	ABK82002	Abk82002 DNA encod												
12	1091	68.1	12923	1	AAN90338	Aan90338 Sequence												
13	1091	68.1	13957	6	ABK81959	Abk81959 cDNA encod												
14	1091	68.1	13957	6	ABT10904	Abt10904 Human bre												
15	1091	68.1	13957	6	ABN95786	Abn95786 Gene #228												
16	1091	68.1	13957	6	ABS69900	Abs69900 Human dys												
17	1091	68.1	13977	6	ABS70403	Abs70403 Human bon												
18	1087.8	67.9	11058	6	AAD37229	Aad37229 Human dys												
19	1085.8	67.8	1991	6	AAD37231	Aad37231 Human dys												
20	932	58.2	4402	3	AAZ48568	Aaz48568 A rod sho												
21	869.2	54.3	4402	3	AAZ48567	Aaz48567 A rod sho												
22	863	53.9	3858	6	AAD37237	Aad37237 Human dys												
23	863	53.9	4825	6	AAD37257	Aad37257 Adeno-ass												

24	863	53.9	4848	6	AAD37263	Aad37263 Adeno-ass
25	863	53.9	5060	6	AAD37264	Aad37264 Adeno-ass
26	854.2	53.4	13815	6	ABK81960	Abk81960 cDNA encod
27	854.2	53.4	13815	6	ABI99799	Abi99799 Mouse isc
28	854.2	53.4	19307	2	AAT27558	Aat27558 Shuttle v
29	854	53.3	5339	6	ABK81998	Abk81998 DNA encod
30	851.2	53.2	3531	6	AAD37238	Aad37238 Human dys
31	851.2	53.2	4498	6	AAD37258	Aad37258 Adeno-ass
32	846.2	52.9	13815	2	AAV18885	Aav18885 Mus muscu
33	843	52.7	3275	1	AAN97129	Aan97129 Partial s
34	766.4	47.9	3999	6	AAD37234	Aad37234 Human dys
35	766.4	47.9	4966	6	AAD37256	Aad37256 Adeno-ass
36	766.4	47.9	4990	6	AAD37262	Aad37262 Adeno-ass
37	765	47.8	1667	6	AAD37235	Aad37235 Human dys
38	732.2	45.7	5462	6	ABK81999	Abk81999 DNA encod
39	706.8	44.1	4075	3	AAZ48569	Aaz48569 A rod sho
40	502.6	31.4	1821	6	AAD37241	Aad37241 Human dys
41	502.6	31.4	2169	6	AAD37232	Aad37232 Human dys
42	501	31.3	1434	6	AAD37243	Aad37243 Human dys
43	450	28.1	1340	6	AAD37239	Aad37239 Human dys
44	409.2	25.6	3747	3	AAZ48566	Aaz48566 A rod sho
45	349.4	21.8	9195	7	ACD19399	Acd19399 cDNA encod

ALIGNMENTS

RESULT 1
AAD37242
ID AAD37242 standard; DNA; 3446 BP.
XX
AC AAD37242;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
(XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 53-54; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

C dystrophy (BMD) in a mammalian subject. The present sequence is human
C dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
C hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and CR
C domain) and 11047-11058 (dystrophin last 3 amino acids)
X
Q Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 U; 0 Other;

Query Match 100.0%; Score 1601; DB 6; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
900 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 959
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 960 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 180
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1020 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 1079
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 CCAGTTTCATACCTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1140 CCAGTTTCATACCTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1199
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1259
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 AACTGAAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1260 AACTGAAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCGAAACTG 480
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1320 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCGAAACTG 1379
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 AAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAACAAAGGAAATGGAGGAAGAG 540
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1380 AAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAACAAAGGAAATGGAGGAAGAG 1439
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 CCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACACATAAGGTGCTTCAA 600
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1440 CCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACACATAAGGTGCTTCAA 1499
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 601 GAAGATCTAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGTT 660
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1500 GAAGATCTAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGTT 1559
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
661 GATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGGA 720
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1560 GATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGGA 1619
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
721 GATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACATC 780
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1620 GATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACATC 1679
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
781 CTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGAA 840
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1680 CTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGAA 1739
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
841 AAAGAGATGCAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTTA 900
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 1740 AAAGAGATGCAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTTA 1799
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
901 TCAAGTCTTCAAAAACCTGGCGTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCATG 960
Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1800 TCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCATG 1859
QY 961 GGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGACC 1020
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1860 GGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGACC 1919
QY 1021 CAGAAACGGAAGCATGGCTGGATAAATTTGCCCCGTTGGGATAATTTAGTCCAAAAA 1080
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1920 CAGAAACGGAAGCATGGCTGGATAAATTTGCCCCGTTGGGATAATTTAGTCCAAAAA 1979
QY 1081 CTTGAAAAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGAT 1140
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1980 CTTGAAAAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGAT 2039
QY 1141 GAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCCGTGGGC 1200
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2040 GAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCCGTGGGC 2099
QY 1201 GATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAA 1260
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2100 GATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAA 2159
QY 1261 ATTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCCGACGCTTACCCT 1320
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2160 ATTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCCGACGCTTACCCT 2219
QY 1321 TTGGCATTCAAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGG 1380
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2220 TTGGCATTCAAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGG 2279
QY 1381 AAGCTTCTGAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGAC 1440
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2280 AAGCTTCTGAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGAC 2339
QY 1441 TTTGTCCAGCATCTCAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCC 1500
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2340 TTTGTCCAGCATCTCAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCC 2399
QY 1501 ATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGGGAC 1560
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2400 ATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGGGAC 2459
QY 1561 CATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCTGACCT 1601
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2460 CATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCTGACCT 2500

RESULT 2
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX AC AAD37260;
XX DT 21-AUG-2002 (first entry)
XX DE Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
PN .WO200183695-A2.
XX XX
PD 08-NOV-2001.
XX XX
PF 27-APR-2001; 2001WO-US013677.
XX XX
PR 28-APR-2000; 2000US-0200777P.
XX XX

(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
Example 1; Page 65-66; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence
Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;
Query Match 99.3%; Score 1590; DB 6; Length 4414;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGGAGCCCAT 60
b 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGGAGCCCAT 1716
Y 61 TCCTTCACAGCAATTTGGAAGCTCCTGGAAGACAAGTCATTTGGCAGTTTCATTGTGAGAG 120
b 1717 TCCTTCACAGCAATTTGGAAGCTCCTGGAAGACAAGTCATTTGGCAGTTTCATTGTGAGAG 1776
Y 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTATCGTGGCTTCTTC 180
b 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTATCGTGGCTTCTTC 1836
Y 181 TGCTGAGGACACATTGCAAGCACAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
b 1837 TGCTGAGGACACATTGCAAGCACAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
Y 241 CCAGTTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
b 1897 CCAGTTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
Y 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAGAAATATCAGAAGATGAAGA 360
b 1957 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAGAAATATCAGAAGATGAAGA 2016
Y 361 AACTGAAGTACAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGTTAGC 420
b 2017 AACTGAAGTACAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGTTAGC 2076
Y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
b 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
Y 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAGAAAGAAAGAAATTTGGAGGAAGA 539
b 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAGAAAGAAAGAAATTTGGAGGAAGA 2196
Y 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACAAACATAGGTGCTTCA 599
b 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACAAACATAGGTGCTTCA 2256
Y 600 AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 659

Db 2257 AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2316
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 779
Db 2377 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 2436
QY 780 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 839
Db 2437 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 2496
QY 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 899
Db 2497 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 2556
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616
QY 960 GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2617 GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2676
QY 1020 CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
Db 2677 CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2736
QY 1080 ACTTGAAGAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGA 1139
Db 2737 ACTTGAAGAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGA 2796
QY 1140 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGG 1199
Db 2797 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGG 2856
QY 1200 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1259
Db 2857 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 2916
QY 1260 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCCGCCAGCTTACCAC 1319
Db 2917 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCCGCCAGCTTACCAC 2976
QY 1320 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATG 1379
Db 2977 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATG 3036
QY 1380 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACACAGGGA 1439
Db 3037 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACACAGGGA 3096
QY 1440 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGC 1499
Db 3097 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGC 3156
QY 1500 CATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACAACTTGTGGGA 1559
Db 3157 CATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACAACTTGTGGGA 3216
QY 1560 CCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
Db 3217 CCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 3258

RESULT 3
ABK81997
ID ABK81997 standard; DNA; 5417 BP.
XX
AC ABK81997;

13-AUG-2002 (first entry)
DNA encoding mini-dystrophin protein deltaR4-R23.
Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
Duchenne's muscular dystrophy; DMD; dystrophin; ds.
Homo sapiens.
Synthetic.
WO200229056-A2.
11-APR-2002.
04-OCT-2001; 2001WO-US031126.
06-OCT-2000; 2000US-0238848P.
(UNMI) UNIV MICHIGAN.
Chamberlain JS, Harper SQ;
WPI; 2002-435334/46.
A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.
Disclosure; Fig 12; 145pp; English.
The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention
Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;
Query Match 88.7%; Score 1420.8; DB 6; Length 5417;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 2; Indels 157; Gaps 2;
1 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 60
1099 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 1158
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTCATGATGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTCATGATGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278
181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTGG 300
1339 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTGG 1398
301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
Qy 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
Qy 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
Qy 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAAGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAAGTGCTTCA 1698
Qy 600 AGAAGATCTAGAAACAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 659
Db 1699 AGAAGATCTAGAAACAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1758
Qy 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1818
Qy 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
Qy 780 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 839
Db 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 1938
Qy 840 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
Qy 900 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
Qy 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
Qy 1020 CCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1079
Db 2119 CCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178
Qy 1080 ACTTGAAAAGAGTACAGCACAGA----- 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTTCACAGGCTGTCAACCACCTCAGCCATCACTAAC 2238
Qy 1103 ----- 1102
Db 2239 ACAGACAACCTGTAATGGAAACAGTAACTACGGTGACCAACAAGGGAACAGATCCTGGTAAA 2298
Qy 1103 -----C 1103
Db 2299 GCATGCTCAAGAGGAACTTCCACCACCCTCCCAAAAAGAGAGGCAGATTACTGTGGA 2358
Qy 1104 CTTTGAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTCGGCCCA 1163
Db 2359 TCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTCGGCCCA 2418
Qy 1164 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTCATTCCTCTCCA 1223
Db 2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTCATTCCTCTCCA 2478
Qy 1224 AGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGT 1283
Db 2479 AGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGT 2538
Qy 1284 GAGCCACGTCAATGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 1343

2539 GAGCCACGTCAAATGACCTTGCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTA 2598
1344 TAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGA 1403
2599 TAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGA 2658
1404 GGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACCT 1463
2659 GGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACCT 2718
1464 TCTTTCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCCTA 1523
2719 TCTTTCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCATCTCGCCAAACAAAAGTGCCCTA 2778
1524 CTATATCAACCACGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTA 1583
2779 CTATATCAACCACGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTA 2838
1584 CCAGTCTTTAGCTGACCT 1601
2839 CCAGTCTTTAGCTGACCT 2856

RESULT 4
AD37240
AAD37240 standard; DNA; 3510 BP.
AAD37240;
21-AUG-2002 (first entry)
Human dystrophin minigene delta3510.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
Homo sapiens.
WQ200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.

Example 1; Page 51-52; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CR domain) and 11047-11058 (dystrophin last 3 amino acids)

XX SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
Query Match 74.5%; Score 1193; DB 6; Length 3510;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGGAGCCCAT 60
Db |||||
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db |||||
Qy 960 TCCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
Qy 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db |||||
Qy 1020 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
Qy 181 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTTAATCATGTGGAAGTGTGTAAGA 240
Db |||||
Qy 1080 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTTAATCATGTGGAAGTGTGTAAGA 1139
Qy 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTGG 300
Db |||||
Qy 1140 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTGG 1199
Qy 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 360
Db |||||
Qy 1200 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 1259
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGCGGAATGCCTCAGGGTAGC 420
Db |||||
Qy 1260 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGCGGAATGCCTCAGGGTAGC 1319
Qy 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAACT 479
Db |||||
Qy 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
Qy 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 539
Db |||||
Qy 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 1439
Qy 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAAACATAAGGTGCTTCA 599
Db |||||
Qy 1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAAACATAAGGTGCTTCA 1499
Qy 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCCTCACATGGTGTGTGTAGT 659
Db |||||
Qy 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCCTCACATGGTGTGTGTAGT 1559
Qy 660 TGATGAATCTAGTGGAGATCACGGCACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 719
Db |||||
Qy 1560 TGATGAATCTAGTGGAGATCACGGCACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 1619
Qy 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTTCCTTTTACAAGACAT 779
Db |||||
Qy 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTTCCTTTTACAAGACAG 1679
Qy 780 CTTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 839
Db |||||
Qy 1680 TTCTGACCAGTGGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 1739
Qy 840 AAAAGAAGATGCAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAATGTT 899
Db |||||
Qy 1740 GAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 1799
Qy 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
Db |||||
Qy 1800 GCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAGAACTAAAGAACCTGTATCAT 1859
Qy 960 GGGCAAACTGTA-----TTCACTCAAAACAAAGATCTTCTTTCAACACTGAAG 1005
Db |||||

1860 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAA 1919
1006 AATAAGTCAGTGACCCAGAGAGCGGAAGCATGGCTGGA----- 1043
1920 ACTCTACGAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACCTGGCT 1979
1044 -----TAACTTTGCCGGTGTGGGATAATTTAGTCCAAAAAACTTGAA----- 1086
1980 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTGGCG 2039
1087 -----AAGAGTACAGCAGACAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCAC 1136
2040 TGA CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCAC 2099
1137 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1196
2100 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2159
1197 GGGCGATCTCCTCATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1256
2160 GGGCGATCTCCTCATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2219
1257 AGAAATTGGCCTCTGAAAGAGAACGCTGAGCCACGTCGAATGACCTTGCTGCGCCAGCTTAC 1316
2220 AGAAATTGGCCTCTGAAAGAGAACGCTGAGCCACGTCGAATGACCTTGCTGCGCCAGCTTAC 2279
1317 CACTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1376
2280 CACTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 2339
1377 ATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAG 1436
2340 ATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAG 2399
1437 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAG 1496
2400 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAG 2459
1497 AGCCATCTGCCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAAACTTGCTG 1556
2460 AGCCATCTGCCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAAACTTGCTG 2519
1557 GGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
2520 GGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2564
RESULT 5
AD37259
ID AAD37259 standard; DNA; 4476 BP.
AC AAD37259;
CT 21-AUG-2002 (first entry)
CX Adeno-associated virus vector plasmid, AAV-MCK-3510.
CX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
CX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
CX Becker muscular dystrophy; ds.
CX Homo sapiens.
CX Unidentified.
CX Chimeric.
CX WO200183695-A2.
CX 08-NOV-2001.
CX 27-APR-2001; 2001WO-US013677.
CX 28-APR-2000; 2000US-020077P.
CX

(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
Example 1; Page 63-65; 7lpp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence
Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;
Query Match 74.5%; Score 1193; DB 6; Length 4476;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 60
Db 1656 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 1715
QY 61 TCCTTCACAGCATTTGGAAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1716 TCCTTCACAGCATTTGGAAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1775
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
Db 1776 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1835
QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1836 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1895
QY 241 CCAGTTTCATACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1896 CCAGTTTCATACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1955
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1956 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2015
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 2016 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2075
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 2076 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2135
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAACACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 539
Db 2136 GAAAGAGTTGAATGACTGGCTAAACAAACACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 2195
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAGGTGCTTCA 599
Db 2196 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAGGTGCTTCA 2255
QY 600 AGAAGATCTAGAACAAGAACACAGTCAAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 659

b 2256 AGAAGATCTAGAACAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 2315
Y 660 TGATGAATCTAGTCGAGATCACGCCAATCTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 719
b 2316 TGATGAATCTAGTCGAGATCACGCCAATCTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 2375
Y 720 AGATCGATGGGAAACATCTGTAGATGACAGAAAGACCGCTGGTCTCTTTTACAAACAT 779
b 2376 AGATCGATGGGAAACATCTGTAGATGACAGAAAGACCGCTGGTCTCTTTTACAAACAG 2435
Y 780 CTTTCTCAAAATGGCAACGCTCTTACTAGAAACAGTGCCTTTTAGTGCAATGGCTTCAGA 839
b 2436 TTCTGACCAGTGGAGCGCTCTGCACCTTTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 2495
Y 840 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT 899
b 2496 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAA 2555
Y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
b 2556 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCAT 2615
Y 960 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTTTCAACACTGAAG 1005
b 2616 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 2675
Y 1006 AATAAGTCAGTGACCCAGAACGGAAGCATGGCTGGA----- 1043
b 2676 ACTCTACGAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 2735
Y 1044 -----TAACTTTGCCCGTGTGGGATAATTTAGTCCAAAACCTTGAA----- 1086
b 2736 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTTGAACCTGCACCTCGC 2795
Y 1087 -----AAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1136
b 2796 TGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 2855
Y 1137 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 1196
b 2856 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 2915
Y 1197 GSGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1256
b 2916 GSGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2975
Y 1257 AGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1316
b 2976 AGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 3035
Y 1317 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1376
b 3036 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 3095
Y 1377 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 1436
b 3096 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 3155
Y 1437 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1496
b 3156 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 3215
Y 1497 AGCCATCTCGCCAAACAAGTGGCCCTACTATATCAACCAGAGACTCAAAACAACCTTGCTG 1556
b 3216 AGCCATCTCGCCAAACAAGTGGCCCTACTATATCAACCAGAGACTCAAAACAACCTTGCTG 3275
Y 1557 GGACCATCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCT 1601
b 3276 GGACCATCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCT 3320

RESULT 6

AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WC200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 43-44; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;

Query Match 68.8%; Score 1102; DB 6; Length 4182;
Best Local Similarity 96.9%; Pred. No. 6.1e-305;
Matches 1134; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACCGAGCCCAT 60
Db |||||
QY 900 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACCGAGCCCAT 959
Db |||||
QY 61 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
Db |||||
QY 960 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1019
Db |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db |||||
QY 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1079
Db |||||
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Db |||||
QY 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
Db |||||
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGTTGG 300

Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGCGCGGTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAATGAAGA 1259
QY 361 AACTGAAAGTACAAGACGAGATGAATCTCCTAAATTTCAGATGGGAATGCCCTCAGGGTAGC 420
Db 1260 AACTGAAAGTACAAGACGAGATGAATCTCCTAAATTTCAGATGGGAATGCCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1379
QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGAGGAAGA 539
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGAGGAAGA 1439
QY 540 GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACATAAAGTGCTTCA 599
Db 1440 GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACATAAAGTGCTTCA 1499
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCACTCACATGGTGGTAGT 659
Db 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCACTCACATGGTGGTAGT 1559
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATGGG 719
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATGGG 1619
QY 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT 779
Db 1620 AGATCGATGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT 1679
QY 780 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTCGCTTTTGTAGTGCATGGCTTCAGA 839
Db 1680 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTCGCTTTTGTAGTGCATGGCTTCAGA 1739
QY 840 AAAAGAAGATGCACTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 899
Db 1740 AAAAGAAGATGCACTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 1799
QY 900 ATCAAGTCTTCAAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1800 ATCAAGTCTTCAAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1859
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAAATAAGTCAGTGAC 1919
QY 1020 CCAGAAGACGGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
Db 1920 CCAGAAGACGGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1979
QY 1080 ACTTGAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1139
Db 1980 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGA 2039
QY 1140 TGAGCTGGACCTCAAGCTGGCCCAAGCTGA 1169
Db 2040 AAAGTTTCTTGGCTGGCTTACAGAAGCTGA 2069

RESULT 7
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX

DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;
Query Match 68.8%; Score 1102; DB 6; Length 5149;
Best Local Similarity 96.9%; Pred. No. 6.8e-305;
Matches 1134; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTATGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTATGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTGCAAGCACCAAGGAGAGATTTCTAATGATGTGCAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTTGCAAGCACCAAGGAGAGATTTCTAATGATGTGCAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1897 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016

Y 361 AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
b 2017 AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
Y 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
b 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2136
Y 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAAGAAAAATGGAGGAAGA 539
D 2137 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAAGAAAAATGGAGGAAGA 2196
Y 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAAACAACATAAGGTGCTTCA 599
b 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAAACAACATAAGGTGCTTCA 2256
Y 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
b 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316
Y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 719
b 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 2376
Y 720 AGATCGATGGGCAAAACATCTGTAGTGCACAGAACCGCTGGGTTCTTTTACAAGACAT 779
b 2377 AGATCGATGGGCAAAACATCTGTAGTGCACAGAACCGCTGGGTTCTTTTACAAGACAT 2436
Y 780 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
b 2437 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 2496
Y 840 AAAAGAAGATGCTAGTGAACAAGATTTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
b 2497 AAAAGAAGATGCTAGTGAACAAGATTTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 2556
Y 900 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
b 2557 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2616
Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
b 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2676
Y 1020 CCAGAAGACGGAGCATGGCTGGATACTTGCCTGGTGGGATAATTTAGTCCAAAA 1079
b 2677 CCAGAAGACGGAGCATGGCTGGATACTTGCCTGGTGGGATAATTTAGTCCAAAA 2736
Y 1080 ACTTGAAGAAGTACAGACAGACACCCCTTGAAAGACTCCAGGAATTCGAAGAGGCCACGGA 1139
b 2737 ACTTGAAGAAGTACAGACAGACACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGA 2796
Y 1140 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGA 1169
b 2797 AAAGTTTCTTGCCTGGCTTACAGAAGCTGA 2826

ESULT 8

AD06794

D AAD06794 standard; DNA; 5952 BP.

C AAD06794;

X 06-AUG-2001 (first entry)

X Human dystrophin gene (Becker form).

X Human; dystrophin; extein; intein; trans-splicing; gene therapy;

W Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.

X Homo sapiens.

S Key Location/Qualifiers

FT misc_recomb 2847..2848
FT /*tag= a
FT /label= S4_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting fragment is
FT used to produce plasmid pSD4"
FT 2952..2953
FT /*tag= b
FT /label= S3_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting fragment is
FT used to produce plasmid pSD3"
FT 3198..3199
FT /*tag= c
FT /label= S2_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting fragment is
FT used to produce plasmid pSD2"
FT 3300..3301
FT /*tag= d
FT /label= S1_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting fragment is
FT used to produce plasmid pSD1"

WO200129243-A1.

26-APR-2001.

13-OCT-2000; 2000WO-CA001216.

15-OCT-1999; 99US-0159868P.

(UYDA-) UNIV DALHOUSIE.

(UYPI-) UNIV PITTSBURGH.

Paul XL, Xiao X;

WPI; 2001-367297/38.

Use of spontaneous or automatic protein splicing to join two or more
peptides at junction site involves expressing extein peptides having co-
reacting portions of split intein attached to them, so that peptides
splice.

Example 1; Fig 2; 81pp; English.

The invention relates to a method directed to the use of spontaneous or
automatic protein trans-splicing to join two or more peptides at junction
site. This method involves expressing extein peptides having co-reacting
portions of split intein attached to them, so that peptides will splice
automatically under suitable conditions. The invention also provides
methods for circumventing virion packaging size limitations in
recombinant virus particle, by splitting a coding region for a protein to
be delivered into two or more extein genes, which are packaged in
separate virus particles and are co-delivered in a target cell for the
expression and for subsequent trans-splicing to form the complete
protein. In particular, the method is used for trans-splicing human
dystrophin and in gene therapies of recombinant adeno-associated virus
(AAV) particles that encode trans-spliced dystrophin, for treating
diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
dystrophy (BMD). The present sequence is human dystrophin gene (Becker
form)

Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;

Query Match 68.2%; Score 1092.4; DB 5; Length 5952;

Best Local Similarity 99.4%; Pred. No. 4.2e-302;

Matches 1107; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60

|||||

Db 891 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 950

2Y 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
2b 951 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1010
2Y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
2b 1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1070
2Y 181 TGCTGAGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
2b 1071 TGCTGAGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130
2Y 241 CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
2b 1131 CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1190
2Y 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGACAGGAAAATTTATCAGAAGATGAAGA 360
2b 1191 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGACAGGAAAATTTATCAGAAGATGAAGA 1250
2Y 361 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCTCAGGGTAGC 420
2b 1251 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCTCAGGGTAGC 1310
2Y 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATC-GAAACT 479
2b 1311 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1370
2Y 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 539
2b 1371 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 1430
2Y 540 GCCTCTTGGACCTGATCTTTGAAGACCTAAACCGCCAAAGTACAACAACATAAAGGTGCTTCA 599
2b 1431 GCCTCTTGGACCTGATCTTTGAAGACCTAAACCGCCAAAGTACAACAACATAAAGGTGCTTCA 1490
2Y 600 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 659
2b 1491 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 1550
2Y 660 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
2b 1551 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610
2Y 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
2b 1611 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1670
2Y 780 CTTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
2b 1671 CTTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1730
2Y 840 AAAAGAAATGCAGTGAACAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
2b 1731 AAAAGAAATGCAGTGAACAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1790
2Y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
2b 1791 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850
2Y 960 GGGCAAACTGTATTCACTCAAAACAAAGATCTTCTTTCAACACTGAAGATAAAGTCAGTGAC 1019
2b 1851 GGGCAAACTGTATTCACTCAAAACAAAGATCTTCTTTCAACACTGAAGATAAAGTCAGTGAC 1910
2Y 1020 CCAGAAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
2b 1911 CCAGAAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1970
2Y 1080 ACTTGAAGAAGGTACAGCACAGACCCCTTGAAAGA 1113
2b 1971 ACTTGAAGAAGGTACAGCACAGGAAACTGAAATA 2004

RESULT 9
ABK82000
ID ABK82000 standard; DNA; 8689 BP.
XX
AC ABK82000;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
XX a nucleic acid sequence encoding the mini-dystrophin peptide.
PS Disclosure; Fig 15; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;
Query Match 68.1%; Score 1091; DB 6; Length 8689;
Best Local Similarity 99.9%; Pred. No. 1.3e-301;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

2Y 540 GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 599
|||
Db 1639 GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1698
|||
2Y 600 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 659
|||
Db 1699 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1758
|||
2Y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATTGGG 719
|||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATTGGG 1818
|||
2Y 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 779
|||
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1878
|||
2Y 780 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
|||
Db 1879 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
|||
2Y 840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
|||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
|||
2Y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
|||
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
|||
2Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1019
|||
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
|||
2Y 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGGGATAAATTTAGTCCAAAA 1079
|||
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGGGATAAATTTAGTCCAAAA 2178
|||
2Y 1080 ACTTGAAAAGATACAGCACAGA 1102
|||
Db 2179 ACTTGAAAAGATACAGCACAGA 2201
|||

RESULT 11
ABK82002
ID ABK82002 standard; DNA; 11443 BP.
XX
AC ABK82002;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX

PT A composition for preparing therapeutic drugs, has a mini-dystrophin

PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Disclosure; Fig 17; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX

SQ Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;

Query Match 68.1%; Score 1091; DB 6; Length 11443;
Best Local Similarity 99.9%; Pred. No. 1.5e-301;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
|||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158
|||
QY 61 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
|||
Db 1159 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
|||
QY 121 TGAAGTAAACCTGACCGCTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
|||
Db 1219 TGAAGTAAACCTGACCGCTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278
|||
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||
Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
|||
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
|||
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGAGAAAATTATCAGAAGATGAAGA 360
|||
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGAGAAAATTATCAGAAGATGAAGA 1458
|||
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
|||
QY 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTTTAATGGATCTCCAGAATC-GAAACT 479
|||
Db 1519 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
|||
QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAGGAAAATGGAGGAAGA 539
|||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAGGAAAATGGAGGAAGA 1638
|||
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 599
|||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 1698
|||
QY 600 AGAAGATCTAGAACAAGAACAAAGTCAGGTCAATTCTCTCACTCACATGGTGGTAGT 659
|||
Db 1699 AGAAGATCTAGAACAAGAACAAAGTCAGGTCAATTCTCTCACTCACATGGTGGTAGT 1758
|||
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATTGGG 719
|||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAAGTATTGGG 1818
|||
QY 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 779
|||
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1878
|||

Y 780 CTTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
|||||
b 1879 CTTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
|||||
Y 840 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
|||||
b 1939 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
|||||
Y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|||||
b 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
|||||
Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 1019
|||||
b 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 2118
|||||
Y 1020 CCAGAAGACGGGAAGCATGGCTGGATAAATCTTCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
|||||
b 2119 CCAGAAGACGGGAAGCATGGCTGGATAAATCTTCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
|||||
Y 1080 ACTTGAAAAGAGTACAGCACAGA 1102
b 2179 ACTTGAAAAGAGTACAGCACAGA 2201
|||||
RESULT 12
AN90338
D AAN90338 standard; cDNA; 12923 BP.
X
C AAN90338;
X
T 29-MAR-1992 (first entry)
X Sequence of human muscular dystrophy (MD) cDNA.
E
W Dystrophin; muscular dystrophy; probe; antibody; diagnosis; prenatal;
W heterozygote; gene therapy; genetic screening; foetal screening; ss.
X
S Homo sapiens.
X
X
H Key Location/Qualifiers
T CDS 209..12923
T /*tag= a
X
N WO8906286-A.
X
X
D 13-JUL-1989.
X
F 16-DEC-1988; 88WO-US004504.
X
R 22-DEC-1987; 87US-00136618.
X
A (CHIL-) CHILDRENS MED CENT.
X Kunkel LM, Monaco A, Hoffman EP, Koenig M;
X
R WPI; 1989-220587/30.
R P-PSDB; AAP90373.
X
T Muscular dystrophy gene - used for prepn. of probes, dystrophic
T polypeptide and antibodies for diagnosis and therapy of muscular
T dystrophy.
X
S Disclosure; Fig 5; 68pp; English.
X
C The inventors claim an MD probe comprising a purified ss NA SQ which
C hybridises to at least a part of the MD gene; pure dystrophin (DS)
C polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The
C probes are equal to or greater than 10b of one of 12 cDNA sequences
C deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd gene
X
Q Sequence 12923 BP; 4296 A; 2613 C; 2963 G; 3044 T; 0 U; 7 Other;

Query Match 68.1%; Score 1091; DB 1; Length 12923;
Best Local Similarity 99.9%; Pred. No. 1.6e-301;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGAGCCCAT 60
|||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCCCTACACGAGCCCAT 1158
|||||
QY 61 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
|||||
Db 1159 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 1218
|||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
|||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1278
|||||
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
|||||
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCCGGTTGG 300
|||||
Db 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCCGGTTGG 1398
|||||
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 360
|||||
Db 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 1458
|||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCCTCAGGGTAGC 1518
|||||
QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAACT 479
|||||
Db 1519 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
|||||
QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 539
|||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 1638
|||||
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAACCCCAAGTACAAACAACATAAGGTGCTTCA 599
|||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACCCCAAGTACAAACAACATAAGGTGCTTCA 1698
|||||
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCAGTGGTGGTAGT 659
|||||
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCAGTGGTGGTAGT 1758
|||||
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
|||||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
|||||
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 779
|||||
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 1878
|||||
QY 780 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
|||||
Db 1879 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
|||||
QY 840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
|||||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
|||||
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|||||
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
|||||
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 1019
|||||
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 2118
|||||

QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAAATTTAGTCCAAAA 2178
QY 1080 ACTTGAAAAAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAAAAGAGTACAGCACAGA 2201
RESULT 13
ABK81959
ID ABK81959 standard; DNA; 13957 BP.
XX
AC ABK81959;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Example 2; Fig 1; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a human dystrophin polynucleotide
CC sequence used in the creation of the mini-dystrophin peptides of the
CC invention
XX
SQ Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;
Query Match 68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.7e-301;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACACATTCGAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTCGAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATCAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATCTCATCAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACCAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACCAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACATAAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACATAAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAACACAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 659
Db 1699 AGAAGATCTAGAACACAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1758
QY 660 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACCAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACCAACTTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1878
QY 780 CTTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 839
Db 1879 CTTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
QY 1080 ACTTGAAAAAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAAAAGAGTACAGCACAGA 2201

RESULT 14
ABT10904
ID ABT10904 standard; cDNA; 13957 BP.
XX
AC ABT10904;
XX
DT 04-DEC-2002 (first entry)

X Human breast cancer associated coding sequence SEQ ID NO: 1038.
X
W Human; breast specific gene; breast cancer; differential expression;
W cytotstatic; gene therapy; gene; ss.
X
S Homo sapiens.

X

N WO200259271-A2.

X

D 01-AUG-2002.

X 25-JAN-2002; 2002WO-US002176.

F 25-JAN-2001; 2001US-0263757P.

R 25-APR-2001; 2001US-0286090P.

R 23-MAY-2001; 2001US-0292517P.

X (GENE-) GENE LOGIC INC.

A

X Orr MS, Nation M, Diggans JC, Zeng W;

I WPI; 2002-674803/72.

R

X Diagnosing breast cancer in a patient comprises detecting the level of
T gene expression in cell or tissue samples, where a differential gene
T expression is indicative of breast cancer.

X Claim 1; SEQ ID NO 1038; 260pp + Sequence Listing; English.

X The present invention relates to methods of diagnosing breast cancer in a
C patient, which comprise detecting the level of expression in a tissue
C sample of two or more genes selected from those shown in ABT09867-
C ABT1112, where a differential expression of the genes indicates breast
C cancer. The methods are useful in diagnosing, treating, detecting the
C progression, and in monitoring treatment of breast cancer in patients.
C The methods are also useful as a screening tool for agents that modulate
C the onset or progression of breast cancer. The breast cancer genes may be
C used as diagnostic markers for the prediction or identification of the
C malignant state of breast tissue, for confirming the type and progression
C of cancer, and for drug screening and assays. The present sequence is a
C coding sequence of the invention. Note: The sequence data for this patent
C did not form part of the printed specification, but was obtained in
C electronic format directly from WIPO at
C ftp.wipo.int/pub.published_pct_sequences

X Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;

Query Match 68.1%; Score 1091; DB 6; Length 13957;

Best Local Similarity 99.9%; Pred. No. 1.7e-301;

Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACAGGAGCCCAT 60

b 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACAGGAGCCCAT 1158

Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120

b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218

Y 121 TGAAGTAAACCTGGACCGTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180

b 1219 TGAAGTAAACCTGGACCGTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Y 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240

b 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

Y 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300

b 1339 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1398

Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360

Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAAACAAGCAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 659
Db 1699 AGAAGATCTAGAAACAAGCAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACATTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACATTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1878
QY 780 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGCTTTCAGA 839
Db 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGCTTTCAGA 1938
QY 840 AAAAGAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 899
Db 1939 AAAAGAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1020 CCAGAACGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTAGTCCAAAA 1079
Db 2119 CCAGAACGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTAGTCCAAAA 2178
QY 1080 ACTTGAAGAAGTACAGCACAGA 1102
Db 2179 ACTTGAAGAAGTACAGCACAGA 2201

RESULT 15

ABN95786

ID ABN95786 standard; DNA; 13957 BP.

XX ABN95786;

XX 13-AUG-2002 (first entry)

XX Gene #2284 used to diagnose liver cancer.

DE Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytotstatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

XX WO200229103-A2.

PN

XX 11-APR-2002.
PD 02-OCT-2001; 2001WO-US030589.
XX 02-OCT-2000; 2000US-0237054P.
PR (GENE-) GENE LOGIC INC.
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX Claim 1; SEQ ID NO 2284; 298pp; English.
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumour in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;
XX Query Match 68.1%; Score 1091; DB 6; Length 13957;
XX Best Local Similarity 99.9%; Pred. No. 1.7e-301;
XX Matches 1102; Conservative .0; Mismatches 0; Indels 1; Gaps 1;
2Y 1 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 60
2Y 1099 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 1158
2Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
2Y 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
2Y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
2Y 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1278
2Y 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
2Y 1279 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
2Y 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCCATCAGGCCGGGTGG 300
2Y 1339 CCAGTTTCTACTCATGAGGGGTACATGATGGATTGACAGCCCCATCAGGCCGGGTGG 1398
2Y 301 TAATATTTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
2Y 1399 TAATATTTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458
2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 420
2Y 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 1518
2Y 421 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
2Y 1519 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578

QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA 599
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1639 GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT 659
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 719
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACGCGCTGGGTTCTTTTACAAGACAT 779
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACGCGCTGGGTTCTTTTACAAGACAT 1878
QY 780 CTTTCTCAATGGCAACGCTCTTTACTGAAGAACAGTCGCTTTTGTAGTCATGGCTTTCAGA 839
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1879 CTTTCTCAATGGCAACGCTCTTTACTGAAGAACAGTCGCTTTTGTAGTCATGGCTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1939 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAA 1079
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAA 2178
QY 1080 ACTTGAAAAGAGTACAGCACAGA 1102
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2179 ACTTGAAAAGAGTACAGCACAGA 2201

Search completed: April 4, 2004, 18:20:56
Job time : 418.121 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: April 4, 2004, 12:04:07 ; Search time 4173.93 Seconds
(without alignments)
16625.161 Million cell updates/sec

itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgcctacacacagg.....taccagttcttagtgacct 1601

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

earched: 3470272 seqs, 21671516995 residues

otal number of hits satisfying chosen parameters: 6940544

linimum DB seq length: 0
laximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase :

GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1420.8	88.7	5417	6	AX538619	AX538619 Sequence
2	1092.4	68.2	5952	6	AR304538	AR304538 Sequence
3	1092.4	68.2	5952	6	AX114289	AX114289 Sequence
4	1091	68.1	8689	6	AX538622	AX538622 Sequence
5	1091	68.1	11443	6	AX538624	AX538624 Sequence
6	1091	68.1	12057	6	AX538627	AX538627 Sequence
7	1091	68.1	13957	6	AX409637	AX409637 Sequence
8	1091	68.1	13957	6	AX538581	AX538581 Sequence
9	1091	68.1	13957	9	HUMDYS	M18533 Homo sapien
10	1091	68.1	13977	6	AR220819	AR220819 Sequence
11	1087.8	67.9	12446	9	HSDMDR	X14298 Human mRNA
12	932	58.2	4402	6	E30220	E30220 Shortened d
13	916.6	57.3	13887	4	AF070485	AF070485 Canis fam
14	869.2	54.3	4402	6	E30219	E30219 Shortened d
15	854.2	53.4	3275	10	MUSDYS	M18025 Mouse dyster
16	854.2	53.4	13815	6	AX306153	AX306153 Sequence
17	854.2	53.4	13815	6	AX538582	AX538582 Sequence
18	854.2	53.4	13815	10	MUSDYSA	M68859 Mouse dyster
19	854.2	53.4	19307	6	AR093392	AR093392 Sequence
20	854.2	53.4	19307	6	AR142592	AR142592 Sequence
21	854	53.3	5339	6	AX538620	AX538620 Sequence
22	732.2	45.7	5462	6	AX538621	AX538621 Sequence
23	706.8	44.1	4075	6	E30221	E30221 Shortened d
24	623.8	39.0	13575	5	GGDYS	X13369 Chicken mRN
25	612.2	38.2	630	9	HSDMDF1	X06179 Human fetal
26	528	33.0	1966	9	HSDMDA1	X06178 Human adult
27	502.6	31.4	1481	9	BC009242	BC009242 Homo sapi
28	409.2	25.6	3747	6	E30218	E30218 Shortened d
29	349.4	21.8	10302	6	AX538583	AX538583 Sequence
30	349.4	21.8	10302	9	HSMUPS	X69086 H.sapiens m
31	340	21.2	10705	10	RNAJ2967	AJ002967 Rattus no
32	334.2	20.9	1898	10	BC024140	BC024140 Mus muscu
33	332.6	20.8	2341	10	BC062163	BC062163 Mus muscu
34	332.6	20.8	2394	10	BC050183	BC050183 Mus muscu
35	332.6	20.8	6045	6	A63605	A63605 Sequence 7
36	332.6	20.8	6045	6	AR281528	AR281528 Sequence
37	332.6	20.8	10320	6	A63607	A63607 Sequence 9
38	332.6	20.8	10320	6	AR281529	AR281529 Sequence
39	327	20.4	327	6	AX538588	AX538588 Sequence
40	326.2	20.4	6059	6	AX107972	AX107972 Sequence
41	326.2	20.4	11096	6	AX538584	AX538584 Sequence
42	326.2	20.4	11096	10	MMY12229	Y12229 M.musculus
43	324	20.2	324	6	AX538613	AX538613 Sequence
44	322	20.1	333	6	AX538590	AX538590 Sequence
45	321	20.0	333	6	AX538589	AX538589 Sequence

ALIGNMENTS

RESULT 1	AX538619	AX538619	5417 bp	DNA	linear	PAT 23-NOV-2002
AX538619	LOCUS	Sequence 39 from Patent WO0229056.				
DEFINITION	AX538619	AX538619				
ACCESSION	AX538619.1	GI:25271163				
VERSION						
KEYWORDS		synthetic construct				
SOURCE		synthetic construct				
ORGANISM		artificial sequences.				
REFERENCE	1					
AUTHORS		Chamberlain, J.S. and Harper, S.Q.				
TITLE		Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL		Patent: WO 0229056-A 39 11-APR-2002;				
		THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

FEATURES		Location/Qualifiers	
source	1..5417	/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="Synthetic"	
RIGIN			
Query Match 88.7%; Score 1420.8; DB 6; Length 5417;			
Best Local Similarity 91.0%; Pred. No. 0;			
Matches 1599; Conservative 0; Mismatches 2; Indels 157; Gaps 2;			
y	1	GAGCTATGCCTACACACAGGCTGCTTATGTTCACCACCTCTGACCCCTACACGGAGCCATT	60
b	1099	GAGCTATGCCTACACACAGGCTGCTTATGTTCACCACCTCTGACCCCTACACGGAGCCATT	1158
y	61	TCCTTCACAGCAATTGGAAGCTCCTGAGACAAAGTCAATTTGSCAGTTTCAATGTATGAGAG	120
b	1159	TCCTTCACAGCAATTGGAAGCTCCTGAGACAAAGTCAATTTGSCAGTTTCAATGTATGAGAG	1218
y	121	TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC	180
b	1219	TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC	1278
y	181	TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA	240
b	1279	TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA	1338
y	241	CCAGTTTCATACCTATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG	300
b	1339	CCAGTTTCATACCTATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG	1398
y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAATAATTATCAGAAGATGAAGA	360
b	1399	TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAATAATTATCAGAAGATGAAGA	1458
y	361	AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTTAGC	420
b	1459	AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTTAGC	1518
y	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATC-GAAACT	479
b	1519	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATCAGAAACT	1578
y	480	GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAAATGGAGGAAGA	539
b	1579	GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAAATGGAGGAAGA	1638
y	540	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAAACAACATAAGGTGCTTCA	599
b	1639	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTACAAACAACATAAGGTGCTTCA	1698
y	600	AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT	659
b	1699	AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT	1758
y	660	TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGGAGAGAACAACTTAAGGTATTGGG	719
b	1759	TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGGAGAGAACAACTTAAGGTATTGGG	1818
y	720	AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGTTCCTTTTACAAGACAT	779
b	1819	AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGTTCCTTTTACAAGACAT	1878
y	780	CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTCTAGTGCATGGCTTTTCTAGA	839
b	1879	CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTCTAGTGCATGGCTTTTCTAGA	1938
y	840	AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	899
b	1939	AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT	1998
y	900	ATCAAGTCTTCAAAAACCTGGCCGTTTAAACCGGATCTAGAAAAAGAAAAAGCAATCCAT	959

Db	1999	ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGGGGATCTAGAAAAAGAAAAAGCAATCCAT	2058
QY	960	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1019
Db	2059	GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	2118
QY	1020	CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA	1079
Db	2119	CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA	2178
QY	1080	ACTTGAAAAAGAGTACAGCACAGA-----C	1102
Db	2179	ACTTGAAAAAGAGTACAGCACAGATTTTCACAGGCTGTACCAACCATCAGCCATCACTAAC	2238
QY	1103	-----C	1102
Db	2239	ACAGACAACTGTAATGGAAAAACAGTAACCTACGGGTGACCACAAGGGAACAGATCCTGGTAAA	2298
QY	1103	-----C	1103
Db	2299	GCATGCTCAAGAGGAACCTCCACCACCACTCCCAAAAGAGGAGGAGATTACTGTGGA	2358
QY	1104	CCTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA	1163
Db	2359	TCCTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA	2418
QY	1164	AGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCA	1223
Db	2419	AGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCA	2478
QY	1224	AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCCTCTGAAAAGAGAAGCT	1283
Db	2479	AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCCTCTGAAAAGAGAAGCT	2538
QY	1284	GAGCCAGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAAGCTCTCACCGTA	1343
Db	2539	GAGCCAGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAAGCTCTCACCGTA	2598
QY	1344	TAACTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTTGCGAGGTGGCCGTCGA	1403
Db	2599	TAACTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTTGCGAGGTGGCCGTCGA	2658
QY	1404	GGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGGACTTTGGTCCAGCATCTCAGCACCT	1463
Db	2659	GGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGGACTTTGGTCCAGCATCTCAGCACCT	2718
QY	1464	TCCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTA	1523
Db	2719	TCCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTA	2778
QY	1524	CTATATCAACCACGAGACTCAAAACAACCTTGTGTGGACCATCCCAAAATGACAGAGCTCTA	1583
Db	2779	CTATATCAACCACGAGACTCAAAACAACCTTGTGTGGACCATCCCAAAATGACAGAGCTCTA	2838
QY	1584	CCAGTCTTTAGCTGACCT	1601
Db	2839	CCAGTCTTTAGCTGACCT	2856
RESULT 2			
AR304538			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
AR304538			
Sequence 1 from patent US 5952.bp			
DNA			
linear			
PAT 12-JUN-2003			
GI:31693691			
Unknown.			
Unknown.			
Unclassified.			
1 (bases 1 to 5952)			
Xiao,X. and Liu,P.X.			
Method and vector for producing and transferring trans-spliced			

JOURNAL Patent: US 6544786-A 1 08-APR-2003; Location/Qualifiers
EATUES source 1. .5952 /organism="unknown" /mol_type="genomic DNA"
peptides
RIGIN
Query Match 68.2%; Score 1092.4; DB 6; Length 5952;
Best Local Similarity 99.4%; Pred. No. 3e-257;
Matches 1107; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
891 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 950
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1010
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 180
1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 1070
181 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
1071 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1130
241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG 300
1131 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG 1190
301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 360
1191 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 1250
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCCTAATGATGTGGAAGTGGTGAAGA 420
1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCCTAATGATGTGGAAGTGGTGAAGA 1310
421 TAGCATGGAAGAAACAAAGCAATTACATAGAGTTTAAATGATCTCCAGAAATC-GAAACT 479
1311 TAGCATGGAAGAAACAAAGCAATTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 1370
480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 539
1371 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 1430
540 GCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAATAAGGTGCTTCA 599
1431 GCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAATAAGGTGCTTCA 1490
600 AGAAGATCTAGAACAAGAACAACTCAGGCTCAATTCCTCACTCAGTGGTGGTGGT 659
1491 AGAAGATCTAGAACAAGAACAACTCAGGCTCAATTCCTCACTCAGTGGTGGTGGT 1550
660 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719
1551 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1610
720 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
1611 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1670
780 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGGCTTTTATGTCATGGCTTTTCA 839
1671 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGGCTTTTATGTCATGGCTTTTCA 1730
840 AAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
1731 AAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790
900 ATCAAGTCTTCAAAAACCTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959

Db 1791 ATCAAGTCTTCAAAAACCTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850
Qy 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACACTGAAGAATAAGTCAGTGAC 1019
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACACTGAAGAATAAGTCAGTGAC 1910
Qy 1020 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAAATTTAGTCCAAA 1079
Db 1911 CCAGAAGACGGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAAATTTAGTCCAAA 1970
Qy 1080 ACTTGAAAAGAGTACAGCACAGCACAGCCCTTGAAAGA 1113
Db 1971 ACTTGAAAAGAGTACAGCACAGCACAGGAAACTGAAATA 2004
RESULT 3
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
Location/Qualifiers
source 1. .5952 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898 /note="S4 junction site"
misc_feature 3198..3199 /note="S2 junction site"
ORIGIN
Query Match 68.2%; Score 1092.4; DB 6; Length 5952;
Best Local Similarity 99.4%; Pred. No. 3e-257;
Matches 1107; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 891 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 950
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1010
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 180
Db 1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 1070
Qy 181 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
Db 1071 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1130
Qy 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG 300
Db 1131 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG 1190
Qy 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 360
Db 1191 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 1250


```
2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 1310

2Y 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT 479
Db 1311 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1370

2Y 480 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAGGAAATGGAGGAAGA 539
Db 1371 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAGGAAATGGAGGAAGA 1430

2Y 540 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1431 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1490

2Y 600 AGAAGATCTAGAAACAAGAACCAAGTCAAGGTCGAATTTCTCTCACTCACATGGTGGTAGT 659
Db 1491 AGAAGATCTAGAAACAAGAACCAAGTCAAGGTCGAATTTCTCTCACTCACATGGTGGTAGT 1550

2Y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 719
Db 1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 1610

2Y 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1611 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1670

2Y 780 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 839
Db 1671 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 1730

2Y 840 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1731 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1790

2Y 900 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
Db 1791 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850

2Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1910

2Y 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1970

2Y 1080 ACTTGAAAAGAGTACAGCACAGACCCCTTGAAAGA 1113
Db 1971 ACTTGAAAAGAGTACAGCACAGGAAACTGAAATA 2004
```

```
RESULT 4
AX538622      AX538622      8689 bp      DNA      linear      PAT 23-NOV-2002
LOCUS        Sequence 42 from Patent WO0229056.
DEFINITION   AX538622
ACCESSION    AX538622
VERSION      AX538622.1  GI:25271171
KEYWORDS     .
SOURCE       synthetic construct
ORGANISM     synthetic construct
              artificial sequences.
REFERENCE    1
AUTHORS      Chamberlain,J.S. and Harper,S.Q.
TITLE        Mini-dystrophin nucleic acid and peptide sequences
JOURNAL      Patent: WO 0229056-A 42 11-APR-2002;
              THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES     location/Qualifiers
              1..8689
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
```

```
ORIGIN
/note="Synthetic"

Query Match      68.1%;   Score 1091;   DB 6;   Length 8689;
Best Local Similarity 99.9%;   Pred. No. 6.8e-257;
Matches 1102;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 1158

QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGTATGAGAG 120
Db 1159 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGTATGAGAG 1218

QY 121 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAGATATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAGATATTATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338

QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 1398

QY 301 TAATATTCTCAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTCAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 1518

QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATC-GAAACT 479
Db 1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578

QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638

QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698

QY 600 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 659
Db 1699 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 1758

QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 1818

QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

QY 780 CTTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 839
Db 1879 CTTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCAGA 1938

QY 840 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1939 AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1998

QY 900 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
Db 1999 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
```

b 2059 GGGCAAACCTGTATTCACCTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
Y 1020 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTTCGGGATAATTTAGTCCAAAA 1079
b 2119 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTTCGGGATAATTTAGTCCAAAA 2178
Y 1080 ACTTGAAAAGAGTACAGCACAGA 1102
b 2179 ACTTGAAAAGAGTACAGCACAGA 2201

ESULT 5
K538624 AX538624 11443 bp DNA linear PAT 23-NOV-2002
OCUS
EFINITION Sequence 44 from Patent WO0229056.
CESSION AX538624
ERSION AX538624.1 GI:25271175
EYWORDS
OURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
EFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES Location/Qualifiers
source 1. 11443
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

RIGIN
Query Match 68.1%; Score 1091; DB 6; Length 11443;
Best Local Similarity 99.9%; Pred. No. 6.8e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCATT 60
b 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCATT 1158

Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGAGAG 120
b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGAGAG 1218

Y 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTCGTGGCTCTTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTCGTGGCTCTTTC 1278

Y 181 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTTGAAAGA 240
b 1279 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTTGAAAGA 1338

Y 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
b 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398

Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
b 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1458

Y 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGAATGCCCTCAGGTTAGC 420
b 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGAATGCCCTCAGGTTAGC 1518

Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATC-GAAACT 479
b 1519 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAACT 1578

Y 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAGGAAATGGAGGAAGA 539
b 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAGGAAATGGAGGAAGA 1638

QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1698

QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACTCATGGTGGTAGT 659
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACTCATGGTGGTAGT 1758

QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1818

QY 720 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 1878

QY 780 CTTTCTCAATGGCAACCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 839
Db 1879 CTTTCTCAATGGCAACCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 1938

QY 840 AAAAGAAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1939 AAAAGAAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998

QY 900 ATCAAGTCTTCAAAAACCTGCCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGCCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058

QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAAGTCAGTGAC 2118

QY 1020 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTTCGGGATAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTTCGGGATAATTTAGTCCAAAA 2178

QY 1080 ACTTGAAAAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 6
AX538627 AX538627 12057 bp DNA linear PAT 23-NOV-2002
LOCUS
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
source 1. 12057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 68.1%; Score 1091; DB 6; Length 12057;
Best Local Similarity 99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 1158

61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
1159 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 1278
181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGTTGG 300
1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGTTGG 1398
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 1578
480 GAAAGAGTTGAAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 539
1579 GAAAGAGTTGAAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 1638
540 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAAATAGGTGCTTCA 599
1639 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAAATAGGTGCTTCA 1698
600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 659
1699 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 1758
660 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719
1759 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1818
720 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
1819 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
780 CCTTCTCAAAATGGCAACGCTCTACTGAAGAACAGTGCCTTTTAAAGATCAAAAATGAAATGTT 839
1879 CCTTCTCAAAATGGCAACGCTCTACTGAAGAACAGTGCCTTTTAAAGATCAAAAATGAAATGTT 1938
840 AAAAGAGATGCAGTGAACAAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
1939 AAAAGAGATGCAGTGAACAAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
900 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
1999 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 2118
1020 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAATTTAGTCCAAAA 1079
2119 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAATTTAGTCCAAAA 2178
1080 ACTTGAAAAGAGTACAGCACAGA 1102
2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 7
AX409637
LOCUS
DEFINITION
Sequence 2284 from Patent WO0229103.
ACCESSION
AX409637
VERSION
AX409637.1 GI:21442342
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE
Gene expression profiles in liver cancer
JOURNAL
Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
source
1..13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
ORIGIN
Query Match 68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGGTGTCTTATGTCAACCACCTCTGACCCCTACCGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGGTGTCTTATGTCAACCACCTCTGACCCCTACCGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAAGTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAAGTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAAGAAAGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAAGAAAGAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 1578
QY 480 GAAAGAGTTGAAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAAATAGGTGCTTCA 599
Db 1639 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAAATAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 659
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719

b 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAGAACAACCTTAAGGTATTGGG 1818
y 720 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGTTCCTTTTACAGACAT 779
b 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGTTCCTTTTACAGACAT 1878
y 780 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTTCAGA 839
b 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTTCAGA 1938
y 840 AAAAGAAGATGCAAGTGAACAAGATTTCACAAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
b 1939 AAAAGAAGATGCAAGTGAACAAGATTTCACAAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
b 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
y 960 GGGCAAACTGTATTCACCTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
b 2059 GGGCAAACTGTATTCACCTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
y 1020 CCAGAAGACGGAAGCATGGCTGGATTAACCTTTCGCCGGTGTGGGATAATTTAGTCCAAAA 1079
b 2119 CCAGAAGACGGAAGCATGGCTGGATTAACCTTTCGCCGGTGTGGGATAATTTAGTCCAAAA 2178
y 1080 ACTTGAAAAGAGTACAGCACAGA 1102
b 2179 ACTTGAAAAGAGTACAGCACAGA 2201

ESULT 8
X538581
OCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
EFINITION Sequence 1 from Patent WO0229056.
CESSION AX538581
ERSION AX538581.1 GI:25271086
EYWORDS Homo sapiens (human)
OURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
EFERENCE Chamberlain, J.S. and Harper, S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 1 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES Location/Qualifiers
source
1. 13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

RIGIN
Query Match 68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
b 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1158
y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGTATGGAGAG 120
b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGTATGGAGAG 1218
y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTC 180
b 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTC 1278
y 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240

Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATACTACTCAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATACTACTCAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGTATTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGTATTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCAGATGGTGGTGTAGT 659
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCAGATGGTGGTGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
QY 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1878
QY 780 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTTCAGA 839
Db 1879 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTTCACCTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTTCACCTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTTGGCCGGTGTGGGATAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTTGGCCGGTGTGGGATAATTTAGTCCAAAA 2178
QY 1080 ACITGAAAAGAGTACAGCACAGA 1102
Db 2179 ACITGAAAAGAGTACAGCACAGA 2201

RESULT 9
HUMDYS
LOCUS HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
DEFINITION Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION M18533 M17154 M18026 M20250
VERSION M18533.1 GI:181856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1699)
Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.
Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
Cell 50 (3), 509-517 (1987)
87273512
87273512
3607877
2 (bases 1678 to 3830)
Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
Conservation of the Duchenne muscular dystrophy gene in mice and humans
Science 238 (4825), 347-350 (1987)
88018015
3659917
3 (bases 1 to 13957)
Koenig,M., Monaco,A.P. and Kunkel,L.M.
The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein
Cell 53 (2), 219-226 (1988)
88194521
3282674
On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.
Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQPSKFGKHIEINFSDL QGRRLLDLLEGLTGOKLPKEKGSTRVHALNVNKAIRVLQNNVDLVNIGSTDIVDG NHKLTGLIWNIIILHWQVKNMKNIMAGLOQTNSEKILLSVWQSTRPQVNVNFT TSWSDGLALNALIHSRFDLFDWNSVVCOSATQRIEHAFNARFYQLGIEKLLDPEDV DTPYDKKSILMYITSLFQVLPQQVSEIAIQEVEMLP RPVKVTKEEHFQLHHQMHYSQ QITVSLAQGYERTSSPKRPFKSYAYTQAAVYVTS DPTSPFPFOHLEAPEDKSFSGSL MESEVNDLRYQTALIEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH QGRVGNILQGLSKLIGTGKLSDEETEVEVQEQMNLNLSRWECLRVASMEKQSNLHRVLM DLQNQKLKELNDWLTKTEERTRKMEEP LGGPDLEDLKRQVQKHVQLQEDLEQEQRVN SILTHMVVVVDESSGDHATAALEEQKVLGDRWANI CRWTEDRWVLIQDILLKWQRLTE EOCILFSAWLSEKEDAVNKIHTTGFKQONEMLSLQKLA VLKADLEKKKQSMGKLYSLK QDILLSTLKNKSVTQTEAWLDNFARCDWNLVQKLEKST AQISQAVTTQPSLTQTVM ETVTTVTITREQIILVHAQEELPPPPQKKGQITVDSEIRKRLDVIDITELHSWITRSEA VLOQSPFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRAQALVEQMVNEGVNA DSIKQASEQNLNRWIEFCQLLSERLNWLEYQNNIIAFYNQLOOLEQMTTAAENWLKIQ PTPPSEPTAKISQLKICKDEVNRLSGLQPQIERLKIQSIALKXKGGQGMFLDADFVAF TNHFKQVFSDVQAREKELQTI FDTLPPMYQETMSAIRTWQOSETKLSIPOLSVTDY EIMEQRLGELQALQSSIQEQSGLYLYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK KLSQVLEHCQKLEEQMKNLRKIQNHIIQTLKKWAAEVDVFLKEWPALGDS EILKKQL KQCRLLVSDIQTIQPSLNSVNEGQKIKNAEAPEFASFARLETTELKELNTQWDHMCQVY ARKEALKGGLEKTVSLQDLSMEHEWMTQABEYLERDFEYKTPDELQKAVEEMKRAK EBAQOKEAKVKLLLTESVNSVIAQAPPAQBALKKELETLTNTYNQWLCTRLNGKCKTLE EYVACWHELLLSYLEKANKWLNVEVEFKLTETENIPGGAEEISEVLDLSLENLMRHSNDP NQIRILAQTLTDGGVNDLELINEELETFNFSRWRELHEEAVRRQKLLLEQSIQSAQETFKS LHLIQESLTFIDKQLAAAYIADKVDAQMPOBAQKIQSDLTSHSIEISLEEMKKNQGKEA

AQRVLSQIDVAQKKLQDVSMKFRLFQKPANFELRLQESKMILDEVXMLPALETKSVE
QEVVQSQNLNHCVNLYKSISEVKSEVEMVIKTGRQIQVQKKQTEPNKELDERVTALKLHY
NELGAKVTERKKQLEKCLKLSRKMKEMNVLTEWLAATDMELTKRSAVEGMP5NLDSE
VANGKATQKEIEKQKVHLKSI TEVGEALKTVLGKKTLEDKLSLLNSWIAVTSRAE
EWNLALLEYQKHMETFDQNDHITKWI IQADTLDESEKKKPOQKEDVLKRLKAE LND
IRPKVDSTRDQAANLMNRGDHCKLVEPQISELNRHFAAISHRIKTKRASIPLKELE
QFNSDIQKLLPLEAEIQOQVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQORITDER
KREEIKIKQOLLQTKHNALKDLRSORRKKALEISHQWYQYKQADDDILKCLDDIEKKL
ASLPEPRDERKIEIDRELQKKKEELNAVRRQAEGLSEDDGAAMAVEPTQIQLSKRWRE
IESKFAQFRRRLNFAQIHTVREETMMVTEDMPLIEISYVPSTYLTETIHTV5QALLEVEQ
LLNAPDLCAKDFEDFLKQEEESLKNIKOSLQOSSCRIDI IHSKTKAALQ5ATPVERVKL
QEALSQDLDFQWEKVNKMKDRQGRFDRSVEKWRPFHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGQRQTVVTLNATGEEIIQSSKTDASILQEKLGSLNL
RWQEVCKQLSDRKKRLEEQNILSEFORDLNEFVLWEADNIIASIPLEPGKQQLKE
KLEQVKLLVEELPLRQGIKQLNETGPPVLVSAPISPEEQDKLENKLTQNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKKLEDLEEQNLHLLWLSPIRNQLEIYNQPNQEGPFD
VQETIAVQAQKQPDVEEILSKGHLYKEKPAQOPVKRKKLEDLSSEWKAVNRLQLQELRA
KQDPLAPGLTTIGASPTQTVTLVTPQVVTKETATISKLEMPSSLMLEVPALADENRAWT
ELTDWLSLLDQVIKSRVMVGDLEDINEMI IKQATMQDLEQRRPQLEELITAAQNLIK
NKTSNQEARTIITDRIERIQNWDEVOEHLQNRQQLNEMLKDSTQWLEAKEEAEQVL
GQAKLESWKEGPYTVDAIQKKTETQKADLRQWQNTVDVANLALKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSEREAALEETHRLLOQFPDLDEKFLAWLTEAETTA
NVLODATRKERLLED5GVKELMKOWDQLOGEIETHADTVYHNLDENSQKILRSLEGS
DAVLLQRRLDNMNFKWSLKKSLNIRSHLEASDDQWKRLHLSLQELLVWLQKDDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTEPEVIMSTLETVRIFLTEQPLEGLEKLYQ
EPRELPEERAQNVTRLLRQAEVNTWEKLNHLSADWQKIDETLERLQELQEA
ELDLKLQAEVIKG5WQVGDLLIDSLQDHLKVKALRGEIAPLKENVSHVNDLARQL
TTLGTLQSPYNLSTLEDLNRWKLLOVAVEDRVQLHEAHRDFGPASQHFSTSVQGP
WERAISPNKVPYYINHETQTTCDWHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
LCLDLLSLSAACDADLQHNLKQNDQPMIDILQIINCLTIIYDRLEQEHNNLVNVPFLCVD
MCLNWLNNVYDTRGRIRVL5FTGIIISLCKAHLDEKRYILFKQVASSSTGFCDDQRRLL
GLLLHDSIQIPROLGEVASF5GGSNIEPSVRSQFANNKPEIEAALFLDMRLEPQSM
VWLPVLHRVAAEATAKHOAKNICKECPIIGFRRYSLKHFNYDICQSCFFSGRVAKGH
KQHYPMVEYCTPTTSGEDVRDFAKVLKNKFRKRYFAKHPRMGYLPVQTVLEGDNMET
PVTLINFPVVD5APASSPOLSHDDTHSRIEHYASRLAEMENSNGSYLND5ISPNE5ID
DEHLIIQHYCQSLNQD5PLSQSPSPAQILISLESEERGELEIRILADLEENRNLOAEY
DRLKQOHEHKGL5PLSPPEMPTSPQSPRDABLIAEAKLLRQHKGRLEARMQIILEDH
NKQLESQHLRLQLLEQPAEAKVNGTTVSSPSTSLQRS5SSQPMLLRVVGSQTSDSM
GEEDLLSPQDTSITGL5EEVMEQLN5FPSSRGRTPGKPMREDTM"

ORIGIN

Query Match	68.1%;	Score	1091;	DB	9;	Length	13957;
Best Local Similarity	99.9%;	Pred.	No. 6.9e-257;				
Matches	1102;	Conservative	0;	Mismatches	0;	Indels	1; Gaps 1;
Qy	1	GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCA	60				
Db	1099	GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCA	1158				
Qy	61	TCCTTCACAGCATTTGGAAGCTCCTGAAGACACAAGTCATTTGGCAGTTCATTGATGGAGAG	120				
Db	1159	TCCTTCACAGCATTTGGAAGCTCCTGAAGACACAAGTCATTTGGCAGTTCATTGATGGAGAG	1218				
Qy	121	TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC	180				
Db	1219	TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC	1278				
Qy	181	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240				
Db	1279	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338				
Qy	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	300				
Db	1339	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	1398				
Qy	301	TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTCAGAAGATGAAGA	360				
Db	1399	TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTCAGAAGATGAAGA	1458				
Qy	361	AACTGAAGTACAAGACAGCATGAATCTCCTAAATTCACATGGGAATGCCTCAGGTTAGC	420				
Db	1459	AACTGAAGTACAAGACAGCATGAATCTCCTAAATTCACATGGGAATGCCTCAGGTTAGC	1518				

```
Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAAACT 479
b 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAACT 1578

Y 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
b 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638

Y 540 GCCTCTTGGACCTGATCTTGAAGACCTFAAAACGCCAAGTACAAACAACATAAAGTGCTTCA 599
b 1639 GCCTCTTGGACCTGATCTTGAAGACCTFAAAACGCCAAGTACAAACAACATAAAGTGCTTCA 1698

Y 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 659
b 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 1758

Y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 719
b 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1818

Y 720 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 779
b 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Y 780 CTTTCTCAAAATGSCAACGCTCTTACTGAGAACAGTGCCTTTTAGTGCAATGGCTTTTTCAGA 839
b 1879 CTTTCTCAAAATGSCAACGCTCTTACTGAGAACAGTGCCTTTTAGTGCAATGGCTTTTTCAGA 1938

Y 840 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
b 1939 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998

Y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 959
b 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058

Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 1019
b 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 2118

Y 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1079
b 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178

Y 1080 ACTTGAAGAAGAGTACAGCACAGA 1102
b 2179 ACTTGAAGAAGAGTACAGCACAGA 2201
```

```
ESULT 10
R220819
OCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
EFINITION Sequence 60 from patent US 6426186.
CCESION AR220819
ERSION AR220819.1 GI:23327696
EYWORDS
OURCE Unknown.
ORGANISM Unknown.
Unclassified.
EFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmath,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES
source
1..13977
/organism="unknown"
/mol_type="genomic DNA"
```

ALIGN

Query Match 68.1%; Score 1091; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCTACACGGAGCCCAATT 1158

QY 61 TCCTTACAGCACTTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 120
Db 1159 TCCTTACAGCACTTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 1218

QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAAAGTATATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAAAGTATATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1398

QY 301 TAATATTCTACAAATTCGGAAGTAAGCTGATTTGGAACAGGAAAAATATCAGAAAGATGAAGA 360
Db 1399 TAATATTCTACAAATTCGGAAGTAAGCTGATTTGGAACAGGAAAAATATCAGAAAGATGAAGA 1458

QY 361 AACTGAAGTACAAGACAGATGAATCTCCTTAAATTCGAAGTGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGACAGATGAATCTCCTTAAATTCGAAGTGGGAATGCCTCAGGGTAGC 1518

QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAACT 1578

QY 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638

QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1698

QY 600 ABRAGATCTAGAACAAAGACAAAGTCAAGGCTCAATTTCTCACTCACTCATGGTGGGTAGT 659
Db 1699 ABRAGATCTAGAACAAAGACAAAGTCAAGGCTCAATTTCTCACTCACTCATGGTGGGTAGT 1758

QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 1878

QY 780 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 839
Db 1879 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 1938

QY 840 AAAAGAAGATGCAGTGAACAAGATTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998

QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058

QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 2118

QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178

QY 1080 ACTTGAAGAAGAGTACAGCACAGA 1102
```


b 1409 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1468
y 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 539
b 1469 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAACAGGAAAAATGGAGGAAGA 1528
y 540 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 599
b 1529 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 1588
y 600 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATCTCTCACTCACATGGTGGTGTAGT 659
b 1589 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATCTCTCACTCACATGGTGGTGTAGT 1648
y 660 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTTGGG 719
b 1649 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTTGGG 1708
y 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAAACAT 779
b 1709 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAAACAT 1768
y 780 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTACTGTCATGGCTTTTCA 839
b 1769 CCTGCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTACTGTCATGGCTTTTCA 1828
y 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 899
b 1829 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1888
y 900 ATCAAGTCTTCAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
b 1889 ATCAAGTCTTCAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1948
y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1019
b 1949 GGGCAAACTGTATTCAATCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2008
y 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
b 2009 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2068
y 1080 ACTTGAAAAGAGTACAGCACAGA 1102
b 2069 ACTTGAAAAGAGTACAGCACAGA 2091

RESULT 12

30220
OCUS E30220 Shortened dystrophin. 4402 bp DNA linear PAT 18-JUN-2001
EPIFINITION E30220 Shortened dystrophin.
CESSION E30220
E30220.1 GI:13017027
EYWORDS JP 199318467-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4402)
AUTHORS Sinichi,T.
TITLE Shortened dystrophin
JOURNAL Patent: JP 199318467-A 3 24-NOV-1999;
SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
OS Unidentified
PN JP 199318467-A/3
PD 24-NOV-1999
PF 08-MAY-1998 JP 1998142134
PR
PI SINICHI TAKEDA
PC C12N15/09,A61K48/00,C12N15/00
CC Strandedness: Both;
CC Topology: Linear;
FH Key Location/Qualifiers

FEATURES
source
1. .4402
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 58.2%; Score 932; DB 6; Length 4402;
Best Local Similarity 79.7%; Pred. No. 7.4e-218;
Matches 1277; Conservative 0; Mismatches 0; Indels 325; Gaps 2;
Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGTACCCCTACACGAGCCCAT 60
Db 1095 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGTACCCCTACACGAGCCCAT 1154
Qy 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCAATTGATGGAGAG 120
Db 1155 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCAATTGATGGAGAG 1214
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTTATCGTGGCTTCTTTC 180
Db 1215 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTTATCGTGGCTTCTTTC 1274
Qy 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1275 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1334
Qy 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1335 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1394
Qy 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAAGATGAAGA 360
Db 1395 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAAGATGAAGA 1454
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1455 AACTGAAGTACAAGAGCAGATGAATCTCCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 1514
Qy 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1515 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1574
Qy 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAATGGAGGAAGA 539
Db 1575 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAATGGAGGAAGA 1634
Qy 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 599
Db 1635 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 1694
Qy 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 659
Db 1695 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 1754
Qy 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTTGGG 719
Db 1755 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTTGGG 1814
Qy 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAT 779
Db 1815 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAT 1874
Qy 780 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTTAGTGCATGGCTTTTCA 839
Db 1875 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTCCCTTTTTAGTGCATGGCTTTTCA 1934
Qy 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1935 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1994

2Y	900	ATCAAGTCTTCAAAAAC	TGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT	959
Db	1995	ATCAAGT	-----	2001
2Y	960	GGGCAAACTGTATTCA	CTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1019
Db	2002	-----	-----	2001
2Y	1020	CCAGAAGACGGAAGCAT	GGCTGGATAAATTTGCCCGGTGTTGGGATAAATTAGTCCAAAA	1079
Db	2002	-----	-----	2001
2Y	1080	ACTTGAAAAGAGTACAG	CACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA	1139
Db	2002	-----	-----	2001
2Y	1140	TGAGCTGGACCTCAAG	CTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG	1199
Db	2002	-----	-----	2001
2Y	1200	CGATCTCCTCATTTGAC	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA	1259
Db	2002	-----	-----CTCGAGAAAGTCAAGGCACCTTCGAGGAGA	2030
2Y	1260	AATTGCGCCTCTGAAAG	GAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	1319
Db	2031	AATTGCGCCTCTGAAAG	GAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	2090
2Y	1320	TTTGGGCATTGAGCTCT	CACCCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATG	1379
Db	2091	TTTGGGCATTGAGCTCT	CACCCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATG	2150
2Y	1380	GAAGCTTCTGAGGTGG	CGCGTCGAGGACCGAGTCAGGAGCTGTCATGAAGCCACAGGGA	1439
Db	2151	GAAGCTTCTGAGGTGG	CGCGTCGAGGACCGAGTCAGGAGCTGTCATGAAGCCACAGGGA	2210
2Y	1440	CTTTGGTCCAGCATCTC	AGCATTCTTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGC	1499
Db	2211	CTTTGGTCCAGCATCTC	AGCATTCTTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGC	2270
2Y	1500	CATCTCGCCCAACAAAG	TGCCCTACTATATCAACCACGAGACTCAAAACAATTGCTGGGA	1559
Db	2271	CATCTCGCCCAACAAAG	TGCCCTACTATATCAACCACGAGACTCAAAACAATTGCTGGGA	2330
2Y	1560	CCATCCCAAAATGACAG	AGCTCTACCAGTCTTTTAGCTGACCT 1601	
Db	2331	CCATCCCAAAATGACAG	AGCTCTACCAGTCTTTTAGCTGACCT 2372	

RESULT 13
AF070485
LOCUS
DEFINITION Canis familiaris dystrophin mRNA, complete cds.
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 13887)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
JOURNAL Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
Location/Qualifiers
1. .13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1. .289
290. .11332
5'UTR
CDS

3'UTR
repeat_region

ORIGIN

```
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLMWEEVEDCYEREDVQKKTFTKWNNAQFSKFGKOHIEINLPSDL
ODGRLDLLLEGLTGQKLPEKKGSTRVHALNNVKALRVLQKNVNDLVNIGSTDIVDG
NHKRLTGLIWNIIILHWQKNVMKNIIMAGLQQTNSEKILLWSVRQSTRNYPQVNVINF
TSWSDGLALNALIHSRDLFDWNSVVCQOSATQRLEHAFNIAKYQLGIEKLLDPEDV
ATTYDPDKSILMYITSLFQVLPPQVSIIEAIQEVMLPRPSQVTRREEHFOIHHQMHYSQ
QITVSLAQGYERAPSPFKPRFKSYAYTAAAYVITSDPTRSPLPSQHLETBEDKSGFRS
LTETEANLDSYQTAALEEVLSWLLSAEDALQAQGISNDVEEVKEQFHTHEGMMDLTS
HQRVGNVVLQGSOLIGTCKLSEDEETEVOEQMNLNSRWECLRVASMEKQSNLHKVL
MDLQNLQKLKELNDWLTTEERTKMEKEPLGPDIEDLKRQVQKHVQLQEDLEQEQVRV
NSLTHMVVVVDESSGDHATAALEEQKVLGDRWANI CRWTEDRWVLLQDILLKWQRF
EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVLSNLQKLAVLKTDLEKKQTKMDKLCSL
NQDILLSALKNTVVAHKMEAWLNDFAQRWDNLVQKLEKSSAQISOAVTTTQPSLTQTTV
METVMVTTREHILVKAQEELPPPPQKQRIIVDSEIRKRLDVIDTELHSHWITRSE
AVLOSPEFAIYRKEGNFSDLKEKVNAIEREKAEPKQLQDASRSAQALVEQMVEGYN
ADSIKQASEQLNSRWIEFQILLSERLNWLEYQNNIITFYNLQQLQEQMTTAEWNLKT
QPTTSEPTAIKSQKICKDEINRLSALQPOIERLKIQSIALKEKGCGPMFLDADFVA
FTNHFNQVFAADVQAREKELQTI FDSLPPMRYQETMSTILTWIQQSETKLSIPQVTVTE
YDIMEQRLGELQALQSSLQEQQNLNLYLSTTVKEMSKKAPLSDISRKYQSEFEIEGR
WKKLSQLVEHCQKLEEQAKLRKIQNHAIKTLKKWITEVDVFLKEEWPALGDSILKR
QLKQCRLLVNDIQTIQPSLNSVNEGACQKMEAEPEFAGRLETELRELNTQWDMCRQ
VYARKEALKGGLDKTVSLQKDLSEHMEWMTQABEEYLERDFEYKTPDELQTAVEEMKR
AKEBAQQKEAKVLLTESVNSVIAQAAPPAQEAALKKELDTLTITNYQWLCITRLNGKCT
LEEYACWHELLSYLEKANKWLSVEVKLKTENISGGAEEIAEVLDSLENLMQHSSE
NPNQIRILAQTLTDGGVNDLINEELETNFSNRRELHBEAVRRQKLLQESIQSAQEI
KSLHLIQESLSSIDKQLAAAYIADKVDAACQPOEAQKIQSDLTSHSIEISLEEMKHNQKG
ETAQRVLSQIDVAQKQLQDVSMKFRLFQKPAPEQRLOQESKMILDEVKHLPALETKS
VEQEVVQSQNLHCVNLKSLSEVKSEVEMVIXTKQVIXQKQVTEPNKELDERVTALKL
HYNELGAKVTERKQLEKCLKLSRMRKEMNALTEWLAATDMELTKRSAVEGMPNSLD
SEVANGKATQKEIEKQVHLKSVTEVGEALKTVLGKKEMLVEDKLSLNSNWIATVSR
AEEWNLILLEYYQKMETFDQNVDIITNWIIQADALLDESEKKPKQOKEDILKRLKAEM
NDIREKVDSTRDQAANLMANRGDCHCKVVEPKISELNHRFAAI SHRIKTGKASIPLKE
LEQNSDIQKLLLEPEAEIQQGVNLKEEDFNKDMSEDNEGTVKELLQRGDNLQQRITD
ERKREHIIKIQOLLQTKHNALKDLRSORRKKALEISHQWYQYKQADDLILKCLDIEK
KLASLPEPRDERKITEIDRELQKKKEELNAVRAQEAEGLEDGGAAMAVEPTQIQLSKRW
REIESKFAQFRRLNFAQIHTVHEESVVAMTDEMFEIISYVPSTYLTETIHTVSOALSEV
EELNAPDLCAQDFEDLQKQESLKNIKOSLQISGRIDIHNKKTAAHLSATPAERA
KLQEALSRLDFQWERNVNNYKDRQGRFDRSVEKWRFRHYDMKILNQWLTEAEQFLKKT
QIPENWEHAKYKWLKELQDGIQQRQSVVRVLNATGEEIIQSSKTASILQEKLGSL
NLRWQEVCKQLAERKKRLEEKNILSEFQRDVNEFVFWLEEDADVANIPLDPGNEQQL
KEKLEQVKLLAEELPLROGILKQNLNETGGTVLVSAPLSPEDQKLENKQTNLQWIK
VSRNLPEKQEEIEAHVKDLGLEEQNLHLLWLSPIRNQLEIYNQPNQTPGPFDIKEIE
VAVQAKQPDVEGILSKGQHYKEKPATQPAKRLLEDLSLSDWKVVTQLQELRAKQPGP
APGLTVPRAQPSQTVTLVTPAVTKETAISKLEMPSSLLLEVPALADFNRAWTETLDW
LSLDRVIKSORVMVGDLIEDINEMI IKQKATQLEQRRPQLEELITAAQNLKKNKTSN
QEARTIITDRIERIQSQWDEVQEHLONRRLQLTMLKDSQWLEAKEEAEQVLQARA
KLESWEAPYTVDAIQKKITETKQAKDLROWQNVNDVANDLALKLRLDYSADDTRKV
HMITENINASWASIHKLSEREALEETHRLLQFPDLLEKFLAWLTEREATTANVLQD
ATHKERLLEDKSGVRELMKQWDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL
QRRLDNMNFKWSELKKSILNIRSHLEASSDQWKRLHLSLOELLVWLQKDDLSRQAP
IGGDFPAVQKQNDVHRAFKRELKTKPEVIMSTLETVRIFLTEQPLEGLEKLYQEPREL
PPEERAQNVTRLRQAEVNTQWEKLNHVSADQWKIDEALERLQELQEAATDELCLK
LRQAEVIKGSWQPVGDLILDSLDHLEKVKALRGEITPLKENVSYVNDLARQLTTLGI
QLSPYNLNTLEDLNRWLKLQVAIEDRIRQLHEAHRDFGPASQHFSLSTVQGPWERAI
SPNKVPYVINHETQTTCDWHPKMTELYQSLADLNNVRFSAYRTAMKRLRLQKALCLDL
LSLSAACDALDQHNLKQNDQPMILQVINCLTITYDRLEQEHNNLVNVLQVDMCLNW
LLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASTGTCFQDRLGLLLH
DSIQIPRQLGEVASFGGSIPEPSVRSFCFQFANNKPEIEAALFLDMRLEPQSMVWLPV
LHRVAAETAQKQAKCNICKECPIIGFRYRSKLFHNYDICIQSCFFSGRVAKGHKMHYP
MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPQVTVLEGNNETPVTLL
NFWPVDAPASSPQLSHDDTHSRIEHYASRLKKMNSNGSYLNDNISPNESIDDEHLL
IQHYWRSNLNQESPLSQSPQAQILISLESEERGELERILADLEGRRNINQAEYDRLKQ
QHEHKGLSPLSPPEMMPPTS PQSPRDAELIAEAKLLRQHKGRLEARMQILEDHKNQOLE
SQLHRLRQLLEQPOAEAKVNGTTVSSPSTLSQRSDSSQPMLLLRVVGSGTSESNGEEDL
LSPPDQSTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM"
11333. .>13887
11789. .11828
/rpt_type=tandem
/rpt_unit="ca"
```


Query Match		57.3%;	Score 916.6;	DB 4;	Length 13887;
Best Local Similarity		90.0%;	Pred. No. 4.8e-214;		
Matches 993;		Conservative 0;	Mismatches 109;	Indels 1;	Gaps 1;
y	1	GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTCTGACCTACACGGAGCCCATTT	60		
b	1183	GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTTCTGACCCACACGGAGCCCACT	1242		
y	61	TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG	120		
b	1243	TCCTTCACAGCATTTTGAAGACTCCTGAAGACAAAGTCATTTGGCCGGTCATTGACAGAGAC	1302		
y	121	TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAAGAGTATTTATCGTGGCTTCTTTC	180		
b	1303	CGAAGCAACCTGGACAGTTATCAACACAGCTTTGAAGAGTACTCTCGTGGCTTCTTTC	1362		
y	181	TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCATATGATGTGGAAGTGGTGAAGA	240		
b	1363	AGCTGAGGATGCACCTGCAAGCCCCAAGGAGAGATTTCATATGATGTGGAAGAGTGAAGA	1422		
y	241	CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGTGG	300		
b	1423	ACAATTTCATACTCATGAGGGATATATGATGGACTTGACATCCCATCAGGGACGGGTGG	1482		
y	301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA	360		
b	1483	TAATGTTCTCCAACCTGGGAAGTCAACTGATTGGAAACAGGGAAATTTATCAGAAGATGAAGA	1542		
y	361	AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC	420		
b	1543	AAACGAAGTGCAGGAACCAAAATGAATCTCCTCAATTCGAAGATGGGAATGCCTCAGGGTAGC	1602		
y	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCAGAAATC-GAAACT	479		
b	1603	TAGCATGGAAAAACAAAGCAATTTACATAAAGTTCTAATGGATCTCAGAAATCAGCAACT	1662		
y	480	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA	539		
b	1663	GAAAGAGTTAAATGACTGGCTAAACAAAAACAGAAAGAGAGAACAAAGGAAAAATGGAGGAAGA	1722		
y	540	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAACAAATAAGGTGCTTCA	599		
b	1723	GCCCTTGGACCTGATATTGAAGACCTTAAACGCCAAGTACAACAAATAAGGTGCTTCA	1782		
y	600	AGAAGATCTAGAACAAAGAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT	659		
b	1783	AGAAGACTTAGAACAGGAACAAAGTCAGGGTCAATTTCCCTCACTCATATGGTGGTGGTAGT	1842		
y	660	TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGG	719		
b	1843	CGATGAATCTAGTGGAGACCATGCAACTGCTGCTTTGAAGAACAACCTTAAGGTACTGGG	1902		
y	720	AGATCGATGGGCAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT	779		
b	1903	AGATCGATGGGCAACATCTGTAGGTGGACAGAAAGATCGCTGGGTTCTTTACAAGACAT	1962		
y	780	CCTTCTCAAAATGGCAACGTTCTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA	839		
b	1963	CCTCCTAAAAATGGCAGCGTTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCGGA	2022		
y	840	AAAAGAAAGTGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT	899		
b	2023	GAAGGAAGATGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAGTGAATGTT	2082		
y	900	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT	959		
b	2083	ATCAATCTTCAGAAACTGGCTGTCTTAAAAACAGATCTGGAAGAGAAAGCAAAACCAT	2142		
y	960	GGGCAAACTGTATTCTACTCAAAACAAAGATCTTCTTTCAACACTGAAGAAATAAGTCAGTGAC	1019		
b	2143	GGACAAACTCTGCTCACTCAACCAAGACCTTCTTTTCAGCGCTGAAAAACACAGTGGTAGC	2202		

QY	1020	CCAGAAAGACGGAAGCATGGCTGGATAA	CTTTGCCCGGTGTTGGGATAATTTAGTCCAAA	1079
Db	2203	CCACAAGATGGAAGCATGGCTGGACAA	CTTTGCCCAGCGCTGGGATAATTTAGTCCAGAA	2262
QY	1080	ACTTGAAAAGAGTACAGCACAGA	1102	
Db	2263	ACTTGAAAAAAGTTCAGCACAGA	2285	
RESULT 14				
E30219				
LOCUS	E30219		4402 bp	DNA linear PAT 18-JUN-2001
DEFINITION	Shortened dystrophin.			
ACCESSION	E30219			
VERSION	E30219.1	GI:13017026		
KEYWORDS	JP 1999318467-A/2.			
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	1 (bases 1 to 4402)			
AUTHORS	Sinichi,T.			
TITLE	Shortened dystrophin			
JOURNAL	Patent: JP 1999318467-A 2 24-NOV-1999;			
COMMENT	SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY			
	OS	Unidentified		
	PN	JP 1999318467-A/2		
	PD	24-NOV-1999		
	PF	08-MAY-1998	JP 1998142134	
	PR			
	PI	SINICHI TAKEDA		
	PC	C12N15/09,A61K48/00,C12N15/00		
	CC	Strandedness: Both;		
	CC	Topology: Linear;		
	FH	Key	Location/Qualifiers	
	FT	source	1. .4402	
	FT		/organism='Unidentified'.	
FEATURES				
	source	Location/Qualifiers		
		1. .4402		
		/organism="unidentified"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:32644"		
ORIGIN				
Query Match		54.3%;	Score 869.2;	DB 6; Length 4402;
Best Local Similarity		77.7%;	Pred. No. 2e-202;	
Matches 1244;		Conservative 0;	Mismatches 33;	Indels 325; Gaps 3;
QY	1	GAGCTATGCCTACACACAGGCTGCTTATGTCA	CCCTCTGACCCCTACACGGAGCCCATTT	60
Db	1095	GAGCTATGCCTACACACAGGCTGCTTATGTCA	CCCTCTGACCCCTACACGGAGCCCATTT	1154
QY	61	TCCTTCACAGCATTTGGAAGCTCCTGAAGACA	AGTCAATTTGGCAGTTCATTGATGGAGAG	120
Db	1155	TCCTTCACAGCATTTGGAAGCTCCTGAAGACA	AGTCAATTTGGCAGTTCATTGATGGAGAG	1214
QY	121	TGAAGTAAACCTGGACCGTTTATCAAAACAG	CTTTAGAGAAGTATTATCGTGGCTTCTTTC	180
Db	1215	TGAAGTAAACCTGGACCGTTTATCAAAACAG	CTTTAGAGAAGTATTATCGTGGCTTCTTTC	1274
QY	181	TGCTGAGGACACATTTGCAAGCACAAAGGAG	AGATTTCTAATGATGTGGAAGTGGTGAAGA	240
Db	1275	TGCTGAGGACACATTTGCAAGCACAAAGGAG	AGATTTCTAATGATGTGGAAGTGGTGAAGA	1334
QY	241	CCAGTTTTCATCTCATGAGGGGTACATGATG	GGATTTTGACAGCCCATCAGGGCCGGTTGG	300
Db	1335	CCAGTTTTCATCTCATGAGGGGTACATGATG	GGATTTTGACAGCCCATCAGGGCCGGTTGG	1394
QY	301	TAATATTCTACAATTTGGAAGTAAGCTGATT	GGAAACAGGAAAAATTTATCAGAAGATGA	360
Db	1395	TAATATTCTACAATTTGGAAGTAAGCTGATT	GGAAACAGGAAAAATTTATCAGAAGATGA	1454
QY	361	AACTGAAGTACAAGACAGATGAATCTCCTAA	ATTTCAAGATGGGAATGCCTCAGGGTAGC	420

1455 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1514
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT 479
1515 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1574
480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAAAGGAAAATGGAGGAAGA 539
1575 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAAAGGAAAATGGAGGAAGA 1634
540 GCCTCTTGGACCTGATCTTGAAGACCTAAAAAGCCAAAGTACAACAACATAAAGGTCTTCA 599
1635 GCCTCTTGGACCTGATCTTGAAGACCTAAAAAGCCAAAGTACAACAACATAAAGGTCTTCA 1694
600 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 659
1695 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1754
660 TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
1755 TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAACAACTTAAGGAGGTCAA 1814
720 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTTACAAGACAT 779
1815 TACTGAGTGGGAAA-----1829
780 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGCTTTCAGA 839
1830 -----1829
840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
1830 -----1829
900 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 959
1830 -----1829
960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
1830 -----ATTGAA 1835
1020 CCAGAAGACGGAAGATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
1836 CTGCACTCCGCTGACTGGC-----1855
1080 ACTTGAAAAGAGTACAGCACAGACAGACCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1139
1856 -----AGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1910
1140 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 1199
1911 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 1970
1200 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1259
1971 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 2030
1260 AATTGGCCCTCTGAAGAGAACGTGAGCCACCTCAATGACCTTGTCTGCCAGCTTACCAC 1319
2031 AATTGGCCCTCTGAAGAGAACGTGAGCCACCTCAATGACCTTGTCTGCCAGCTTACCAC 2090
1320 TTTGGGCATTCAGCTCTCACCGGTATAACCTCAGCACTCTTGGAGACCTGAAACACCATG 1379
2091 TTTGGGCATTCAGCTCTCACCGGTATAACCTCAGCACTCTTGGAGACCTGAAACACCATG 2150
1380 GAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCTGAAGCCCCACAGGGA 1439
2151 GAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCTGAAGCCCCACAGGGA 2210
1440 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC 1499
2211 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC 2270

QY 1500 CATCTCGCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTGCTGGGA 1559
Db 2271 CATCTCGCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTGCTGGGA 2330
QY 1560 CCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCT 1601
Db 2331 CCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCT 2372
RESULT 15
MUSDYS Mouse dystrophin mRNA, partial cds. linear ROD 27-APR-1993
LOCUS M18025.1 GI:192971
DEFINITION dystrophin.
ACCESSION M18025
VERSION 1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 3275)
AUTHORS Hoffmann, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
COMMENT Original source text: Mouse adult heart, cDNA to mRNA.
FEATURES
Location/Qualifiers
1..3275
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
<1..>3275
/note="dystrophin"
/codon_start=3
/protein_id="AAA37530.1"
/db_xref="GI:192972"
/translation="SYAFTQAAYVATSDSTQSPYPSQHLEAPRDKSLDSSLMETEVLNLSYQTALEEVLSWLLSAEDTLRAQGEISNDVEVEKQFHAHEGFMDLTSHQGLVGNVLOLGSQLVGKGLSEDEAEVQEQMNLNSRWECVLRVASMESKQSKLHKVLMDLQNKQLKELHDWLTKEERTKMEEPFDDLEDLKQVQVQKHVQLQEDLEQEQVRVNSLTHMVFVDESSGDHATAALEEQKLVGLDRWANICRWTEDEWTVLQDILLKWOHFTTEOCLFSTWLSEKEDAMKNIQSTGFQDQNMNSSLHKISTLKIDLEKKKPTMEKLSLNQDLISALKNKSVTKMEIWMENFAQRWDLTKLEKSSAQISOAVTTTPFLTQITVTMETVMTVTREQIMVHAQEELPPPPQKKQITVDSELKRDLVDITELHSWITRSEAVLQSSEFAYVRKEGNIQSLQEKVNAIAREKAEKFRKLQDASRAQALVEQMANEGVNAESIRQASEQLNSRWTFEPCQLLSERVNWLEYQNTMTTFYNQQLQEQMTTAEENLLKTQSTLSEPTAISKQLKICKDEVNRLSALQPIBQIKIQLKQLEKQGPFLDADDFVAFTHFNPHITAIKRAKEBELQITFDLTPPMRYQETMSSIRTIWIQQSEKSLVPYLSVTEYEIMEERLFDGVRAKEBELQITFDLTPPMRYQETMSSIRTIWIQQSEKSLVPYLSVTEYEIMEERLGLQALQSSLKEQONGFNYSLDTVKEMAEVDVFLKEWPALGDABEILKKQLKQCRLLVESCQKLEEHMKNLRFQNHITLQKMAEVDVFLKEWPALGDABEILKKQLKQCRLLVGDITQIOPSLSVNEGQKIKSEAELEFASRLTELRNLTQWDHICROVYTRKBAKAGLDKTVSLQKDLSEMHFWMTQAESEYLERDFEYKTPDELQTAVEEMKRAKEALQKE TKVKLLTETVNSVIAHAPPSAQEALKKELETLTTNYQWLCTRLNGKCKTLEEVWACWH ELLSYLEKANKWLNVEVELKLTMTENVDPAGEEITTEVLESLNLMHSEENPNQIRLLA QTLTDGGVMDLINEELETFNSRWRELHEEAVRKQKLLLEQSIQSAQEIEKSLHLIQES LEF"

ORIGIN 744 bp upstream of PstI site.

Query Match 53.4%; Score 854.2; DB 10; Length 3275;
Best Local Similarity 86.5%; Pred. No. 9.5e-199;
Matches 954; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGAGCCCAT 60
Db 2 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCCTA 61
QY 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAACTCATTTGGCAGTTCATTGATGGAGAG 120
Db 62 TCCTTCACAGCATTGGAAGCTCCTGAAGACAACTCATTTGACAGTTCATTGATGGAGAG 121

121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCCTTC 180
122 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAAGTACTTTCATGGCTTCCTTC 181
181 TGCTGAGGACACATTCGAAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
182 TGCCGAGGATACATTGGAGAGCACAAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 241
241 CCAGTTTCTATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
242 ACAGTTTCTATGCTCATGAGGGATTCTATGATGGATCTGACATCTCATCAAGGACTTGTGG 301
301 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
302 TAAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAAATTATCAGAAGATGAAGA 361
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
362 AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC 421
421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAACT 479
422 TAGCATGGAACAAACAAAGCAAAATTACACAAAGTTCTAATGGATCTCCAGAATCAGAAAT 481
480 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAACAAAGGAAAATGGAGGAAGA 539
482 AAAAGAACTACATGACTGTTAAACAAAACCTGAAGAGAGAACTAAGAAAATGGAGGAAGA 541
540 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACACATAAGGTGCTTCA 599
542 GCCCTTTGGACCTGATCTTTGAAGATCTAAATGGCAAGTACAACACATAAGGTGCTTCA 601
600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 659
602 AGAAGATCTAGAACAGGAGCAGGTGAGGTCACTCGCTCACTCACATGGTAGTGGT 661
660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGG 719
662 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGTACTGG 721
720 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT 779
722 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT 781
780 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 839
782 TCTTCTAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTATGATACATGGCTTTCAGA 841
840 AAAAGAAAGATGCAGTGAACAAAGATTCAACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
842 AAAAGAAAGATGCAATGAAGAACATTCAGACAAAGTGGCTTTTAAAGATCAAAATGAAATGAT 901
900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
902 GTCAAGTCTTCAAAAATATCTACTTTAAAATAGATCTAGAAAAGAAAAGCCCAACCAT 961
960 GGGCAAACTGTATTCACTCAACAAAGATCTTCTTCAACACACTGAAGAAATAAGTCAGTGAC 1019
962 GGAATAAATAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAATAAGTCAGTGAC 1021
1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAA 1079
1022 TCAAAAGATGGAATCTGGATGGAATACTTTTGCAACAGTGGGACATTTAAACCCAAA 1081
1080 ACTTGAAAAGAGTACAGCACAGA 1102
1082 ACTTGAAAAGAGTTCAGCACAAA 1104

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

4 nucleic - nucleic search, using sw model
in on: April 4, 2004, 16:52:43 ; Search time 3673.37 Seconds
(without alignments)
17079.796 Million cell updates/sec

title: US-09-845-416-12_COPY_900_3000
affect score: 2101
sequence: 1 gagctatgcctacacacagg.....ttctgcattgattctatccaa 2101

oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
eached: 27513289 seqs, 14931090276 residues
otal number of hits satisfying chosen parameters: 55026578

inimum DB seq length: 0
aximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	787	37.5	3870	11 BC036103	BC036103 Homo sapi
2	753.8	35.9	5691	29 AY399453	AY399453 Homo sapi
3	629.8	30.0	3056	11 AK044536	AK044536 Mus muscu
4	601.8	28.6	5697	29 AY399455	AY399455 Mus muscu

5	509.2	24.2	3753	11 AK081426	AK081426 Mus muscu
6	509	24.2	595	14 CB177816	CB177816 is21c01.x
7	505	24.0	777	14 CD653550	CD653550 AGENCOURT
8	505	24.0	824	9 AL556247	AL556247 AL556247
9	503.4	24.0	801	14 CB991394	CB991394 AGENCOURT
10	503.4	24.0	1098	13 BX365572	BX365572 BX365572
11	494	23.5	620	13 BQ640063	BQ640063 he23g04.Y
12	494	23.5	797	14 CB960722	CB960722 AGENCOURT
13	484.6	23.1	2874	29 AY408546	AY408546 Homo sapi
14	483.8	23.0	508	14 CD701871	CD701871 EST18395
15	480	22.8	728	14 CB228986	CB228986 AGENCOURT
16	471.4	22.4	770	12 BG719710	BG719710 602690430
17	471.2	22.4	2874	29 AY408548	AY408548 Mus muscu
18	460.2	21.9	1047	14 CB850319	CB850319 MRA-0070
19	455.8	21.7	591	13 BX490860	BX490860 DKFZp686F
20	453.8	21.6	652	10 BB629984	BB629984 BB629984
21	453.8	21.6	1298	11 AK087829	AK087829 Mus muscu
22	453.8	21.6	2135	11 AK013510	AK013510 Mus muscu
23	453.8	21.6	4437	11 AK036936	AK036936 Mus muscu
24	452.6	21.5	5676	29 AY399454	AY399454 Pan trogl
25	452.2	21.5	1384	11 AK075809	AK075809 Mus muscu
26	449	21.4	704	10 BB610411	BB610411 BB610411
27	439	20.9	578	13 BX485574	BX485574 DKFZp686L
28	428.6	20.4	854	9 AI196693	AI196693 ui53e10.Y
29	423	20.1	665	13 BY742604	BY742604 BY742604
30	421	20.0	520	29 CG606443	CG606443 OST284316
31	420.2	20.0	512	14 CF162938	CF162938 B0718G07-
32	410.4	19.5	599	10 BB666688	BB666688 BB666688
33	407.2	19.4	554	14 CB613696	CB613696 AMGNNUC:N
34	404.6	19.3	3051	11 BC036095	BC036095 Homo sapi
35	404.4	19.2	493	14 CA888041	CA888041 B0142C06-
36	400.6	19.1	513	29 CG605968	CG605968 OST283399
37	399.6	19.0	508	29 CG606947	CG606947 OST285335
38	398.2	19.0	717	14 CB527785	CB527785 UI-M-FY0-
39	397.4	18.9	495	14 CA894775	CA894775 B0187G06-
40	392.2	18.7	650	13 BY714491	BY714491 BY714491
41	388.6	18.5	488	29 CG597965	CG597965 OST261937
42	388.6	18.5	542	29 CG594848	CG594848 OST253426
43	388.6	18.5	644	13 BU313510	BU313510 603540290
44	383.4	18.2	483	14 CA893902	CA893902 B0182B01-
45	381.8	18.2	423	9 AA460476	AA460476 zx61e10.Y

ALIGNMENTS

RESULT 1
BC036103
LOCUS BC036103 3870 bp mRNA linear HTC 19-NOV-2003
DEFINITION Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (cDNA clone IMAGE:5274415), with apparent retained intron.
ACCESSION BC036103
VERSION BC036103.1 GI:23271310
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3870)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3870)
Strausberg,R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.

Location/Qualifiers
1. .3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 37.5%; Score 787; DB 11; Length 3870;
Best Local Similarity 94.8%; Pred. No. 1.6e-172;
Matches 814; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 60
|||||
1043 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 1102
|||||
61 TCCTTCACAGCATTGGAGCTCCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG 120
|||||
1103 TCCTTCACAGCATTGGAGCTCCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG 1162
|||||
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
|||||
1163 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1222
|||||
181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
|||||
1223 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1282
|||||
241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
|||||
1283 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1342
|||||

301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
|||||
1343 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1402
|||||
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
|||||
1403 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1462
|||||
421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
1463 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1522
|||||
481 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 540
|||||
1523 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA 1582
|||||
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
|||||
1583 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1642
|||||
601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
|||||
1643 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1702
|||||
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGG 720
|||||
1703 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGG 1762
|||||
721 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACAG 780
|||||
1763 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACAT 1822
|||||
781 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAAACTTCTGGTGGGTACAGCT 840
|||||
1823 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTGATGGCTTTTCAGA 1882
|||||
841 GAAAGATCATGAATTAAGC 859
|||||
1883 AAAAGAGATGCAGTGAAC 1901
|||||

RESULT 2
AY399453
LOCUS
DEFINITION
AY399453
Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
5691 bp DNA linear GSS 12-DEC-2003
ACCESSION
AY399453
VERSION
AY399453.1
GI:39755442
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 5691)
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 5691)
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES	Location/Qualifiers	
source	1..5691	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
gene	/db_xref="taxon:9606"	
	<1..>5691	
ICIN	/locus_tag="HCM0229"	
Query Match	35.9%; Score 753.8; DB 29; Length 5691;	
Best Local Similarity	97.4%; Pred. No. 9.6e-165;	
Matches 780; Conservative	0; Mismatches 12; Indels 9; Gaps 1;	
1	GAGCTATGCCTACACACAGCGTGTCTTATGTCTACACACCTCTGTACCCCTACACGGAGCCCAT	60
294	GAGCTATGCCTACACACAGCGTGTCTTATGTCTACACACCTCTGTACCCCTACACGGAGCCCAT	353
61	TCCTTCACAGCATTTGGAAGTCTCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	120
354	TCCTTCACAGC-----TCCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	404
121	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	180
405	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	464
181	TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA	240
465	TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA	524
241	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG	300
525	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG	584
301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA	360
585	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA	644
361	AACTGAGTACAAGACGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC	420
645	AACTGAGTACAAGACGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC	704
421	TAGCATGGAAAAACAAGCAATTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	480
705	TAGCATGGAAAAACAAGCAATTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	764
481	GAAAGATTGAATGACTGGCTTAACAAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA	540
765	GAAAGATTGAATGACTGGCTTAACAAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA	824
541	GCCTCTTGACCTGATCTTTGAGACCTAAACCGCCAAGTACAACAACATAAGGTGCTTCA	600
825	GCCTCTTGACCTGATCTTTGAGACCTAAACCGCCAAGTACAACAACATAAGGTGCTTCA	884
601	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT	660
885	AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT	944
661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	720
945	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	1004
721	AGATCGATGGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAAGACAG	780
1005	AGATCGATGGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAAGACAT	1064
781	TTCTGACCACTGGAAGCGTCT 801	
1065	CCTTCTCAATGGCAACGTCT 1085	
35ULT 3		
044536		
0CUS		
	3056 bp	mRNA linear
	AK044536	HTC 20-SEP-2003

DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930019P21 product:dystrophin, muscular dystrophy, full insert sequence.
ACCESSION	AK044536
VERSION	AK044536.1 GI:26090404
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci,P. and Hayashizaki,Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3056)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES
source
1. .3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930019F21"
/db_xref="MGI:2409854"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. .3056
/note="dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)"

ORIGIN
Query Match 30.0%; Score 629.8; DB 11; Length 3056;
Best Local Similarity 83.4%; Pred. No. 7.2e-136;
Matches 715; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
2Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGGAGCCCAT 60
|||
944 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCTA 1003
|||
2Y 61 TCCTTCACAGCATTGGGAAGTCCTGAAGACAGCAAGTCATTTGGCAGTTCATTGATGGAGAG 120
|||
2Y 1004 TCCTTCACAGCATTGGGAAGTCCTCCAGAGACAAAGTCACCTTGACAGTTCATTGATGGAGAC 1063
|||
2Y 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 180
|||
2Y 1064 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAGAAGTACTTTTCATGGCTTCTTC 1123
|||
2Y 181 TGCTGAGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
|||
2Y 1124 TGCCGAGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 1183
|||
2Y 241 CCAGTTTCATCTCATGAGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG 300
|||
2Y 1184 ACAGTTTCATGCTCATGAGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 1243
|||
2Y 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
|||
2Y 1244 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAATTTATCAGAAGATGAAGA 1303
|||
2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCTCAGGGTAGC 420
|||
2Y 1304 AGCTGAAGTCAAGAACAAATGAATCTCCTAAATCAAGATGGGAATGCTCAGGGTAGC 1363
|||
2Y 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||
2Y 1364 TAGCATGGAACAAACAAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAATCAGAAAT 1423
|||
2Y 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAATGGAGGAAGA 540
|||
2Y 1424 AAAAGAACTAGATGACTGGTTAAACAAAACCTGAAGAGAGAACTAAGAAAATGGAGGAAGA 1483
|||
2Y 541 GCCTCTTGGACCTGATCTTTGAGACCTTAAACGCCAAGTACAACAACTAAGGTGCTTCA 600
|||
2Y 1484 GCCCTTTGGACCTGATCTTTGAGATCTAAATGCAAGTACAACAACTAAGGTGCTTCA 1543
|||
2Y 601 AGAAGATCTAGAACAAAGAACTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
|||
2Y 1544 AGAAGATCTAGAACAGGAGCAGGTTCAGGGTCAACTCGCTCACTCACATGGTGTAGTGGT 1603
|||
2Y 661 TGATGAATCTAGTGGAGATCACCGAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
|||

Db 1604 TGATGATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 1663
|||
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAG 780
|||
Db 1664 AGATCGATGGGCAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT 1723
|||
QY 781 TTCTGACCAGTGGGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGTGTGGTACAGCT 840
|||
Db 1724 TCTTCTAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTACATGGCTTTCAGA 1783
|||
QY 841 GAAAGATGATGAATTAA 857
|||
Db 1784 AAAAGAAGATGCAATGA 1800
|||
RESULT 4
AY399455 5697 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399455
VERSION AY399455.1 GI:39755444
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 5697)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5697)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .5697
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1. .>5697
/locus_tag="HCM0229"

ORIGIN
Query Match 28.6%; Score 601.8; DB 29; Length 5697;
Best Local Similarity 85.5%; Pred. No. 2.8e-129;
Matches 685; Conservative 0; Mismatches 107; Indels 9; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGAGCCCAT 60
|||
Db 300 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACTCTGATTCACACAGAGCCCTA 359
|||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
|||
Db 360 TCCTTCACAG-----GCTCCAGAGACAAGTCACTTGACAGTTCATTGATGGAGAC 410
|||
QY 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
|||
Db 411 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAGAAGTACTTTTCATGGCTTCTTTC 470
|||
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCATATGATGTGGAAGTGGTGAAGA 240
|||

```
|||||
471 TGCCGAGGATACATTGCGAGCACAAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 530
/
241 CCAGTTTCATACATCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCGGGTTGG 300
|||||
531 ACAGTTTCATGCTCATGAGGGATTATCATGATGGATCTGACATCTCATCAAGCACTTGTGG 590
/
301 TAATATTCACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
|||||
591 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAATAATTATCAGAAGATGAAGA 650
/
361 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 420
|||||
651 AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 710
/
421 TAGCATGGAATAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
|||||
711 TAGCATGGAATAAAACAAAGCAATTACACAAAGTTCTTAATGGATCTCCAGAAATCAGAAATT 770
/
481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
|||||
771 AAAAGAACTAGATGACTGGTTTAACAAAAAAGTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 830
/
541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCCAGTACAAACAACATAAGGTGCTTCA 600
|||||
831 GCCCTTTGGACCTGATCTTGAAGATCTAAATGCCCAAGTACAAACAACATAAGGTGCTTCA 890
/
601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
|||||
891 AGAAGATCTAGAACAGGAGCAGGTCAAGGTCAACTCGCTCACTCACATGGTGTAGTGGT 950
/
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGTATTGGG 720
|||||
951 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGGGAAGAACAACTTAAGTACTGGG 1010
/
721 AGATCGATGGCAAAACATCTGTAGTGGACAGAACCGCTGGGTTCTTTTACAAGACAG 780
|||||
1011 AGATCGATGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT 1070
/
781 TTCTGACCAGTGGGAAGCGTCT 801
|||||
1071 TCTTCTAAATGGCAGCATTT 1091
|||||

3RESULT 5
<081426
>CUCS
3FINITION
AK081426 3753 bp mRNA linear HTC 20-SEP-2003
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
library, clone:Cl30016K19 product:dystrophin related protein 2,
full insert sequence.
<CESSION
ERSION
EYWORDS
DURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
3REFERENCE
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
3REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
```

```

Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3753)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramatsu,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
1..3753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:Cl30016K19"
/db_xref="MGI:2413800"
/db_xref="taxon:10090"
/clone="Cl30016K19"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
181..3054
/note="unnamed protein product; dystrophin related protein
2 (MGD|MGI:107432, GB|U43520, evidence: BLASTN, 99%,
match=807)
putative"
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"
/translation="MQPLVMQGCPTLPRCHEWHAADPFHSSSLRNTCPQPVRAAV
TIPAPPWDGADPCLSPKLLNGTVGATGPLEPSAMNLCWNEIKKSHNLRLEAFSD
```

LSGKLQPLRLBIIDWLSQXDEELSAQLPQGDVALVQOEKETHAAFMEEVKSXGPGYIS
SVLESAQAFLSQHPFEELEESQSESKDTSRQRIQNLSRFVWKQATVASSEWEKLTAR
CVDQHRHIEHTLEHLLLEIOGAMEELSLTLQAEGVRAWPEIGDLFIDSLPEHIQAIK
LFKEEFSVXKDVKLNDLAHQLAISDVHLSMENSRALEQINIRWKQLQVSAERLKQ
LQDAHRDFGSOHFLSTSVQVPWERAFSPNKVPYIYNHQAOQTTCDWHPKMTELYOTL
ADLNNIKFSAYRTAMKLRVQKALRLDLVLTTLTALEIFNEHDLQASEHVMADVVEIHC
LTALYERLEERGILVNVPLCVDSINLWLLNVFDSGRSGKMRALSFKTGTACLCGTEV
KEKQLYLSQVANGSQCDQRHLGALLHEAIQVPRQJGEVAAFGGSNVESVRSRFRF
STGKPVIEASOFLEWNVNLEPQSMVWLAVLHRVTIAEOVKHQTKCSICROCPKGFPRYR
SLKQFNVDICOTCFLTGRASKGNKLHYPIMEYVTPHTSSNMRFATTLKNNKFRSKQY
FSKHQFQGYLVPQSVLESDCSETPASSPMLPHADTTHSRIEHPFASLAEEMSONCSFFN
DLSLPPDDSIDEDQYLLRHSSPITDREPAFGQQAPCSMATESKGELEKILAHLEDENRI
LQGE LRRLKWOHEEAEEAPTLLVEGSAEATPDHRNEELAEARILRQHKSRLETRMQIL
EDHNKQLESQIQRRLRELLLQPPSESDGNGSAGSSLASPRQSEGSHPREKQGTTPDTE
VADDVGSKSQVSLCLEDIMEKLRHAFPSVRSDDVTANTILLAS"
3731. .3736
/note="putative"
polyA_signal
3753
/note="putative"
polyA_site

IRIGIN

Query Match		24.2%;	Score 509.2;	DB 11;	Length 3753;
Best Local Similarity		61.1%;	Pred. No. 1e-107;		
Matches 823;		Conservative	0;	Mismatches 523;	Indels 0;
		Gaps	0;		
iy	756	ACCGCTGGGTTCTTTTACAAGACAGATTCTTGACCAGTGGAAAGCGTCTGACACCTTTCTCTGTC	815		
ib	455	ACCTCGCGCTCGCCTAGAGGCTTCTTCAGACCTCAGTGGAAAACCTTCAGCTCCCGCTCC	514		
iy	816	AGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTG	875		
ib	515	GAGAGATTATTGACTGGCTCAGCCAAAAGGATGAGGAGTTGTCAAGCTCAGCTCCCTTGC	574		
iy	876	GAGGCGACTTTCAGCAGTTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGAAT	935		
ib	575	AAGGGATGTGGCCCTGGTACAAACAGGAGAAGGAGACACATGCAGCCTTCATGGAAGAAG	634		
iy	936	TGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAG	995		
ib	635	TCAAGTCTAAGGGCCCTACATCTCTCTGTGCTGTAATCGGCCAGGCTTTTCTGTCCC	694		
iy	996	AGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGACAGAGACTGCCTCCTGAGG	1055		
ib	695	AGCATCCATTGAAGAATTAGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCCAGAC	754		
iy	1056	AGAGAGCCCGAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGT	1115		
ib	755	AGCGGATTCAAACCTTAGTCGCTTTGTATGGAAAGCAGGCAACAGTGGCCAGTGAGTGT	814		
iy	1116	GGGAAAAATTGAACCTGCCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAA	1175		
ib	815	GGGAGAAGCTGACAGCCCGCTGTGTAGATCAGCATCGCCACATTGAGCATACTCTGGAAC	874		
iy	1176	GACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGG	1235		
ib	875	ATCTATTGGAGATCCAAGGGGCAATGGAGGAATTGAGCAGTACTTTGACCCCAAGCAGAGG	934		
iy	1236	TGATCAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCAC	1295		
ib	935	GAGTCCGAGCCACGTGGGAGCCCATAGGAGATCTCTTTATTGATTCCCTCCAGAGCATA	994		
iy	1296	TCGAGAAAGTCAAGGCATTCGAGGAGAAATTGGCGCTCTGAAAGAGAAACGTGAGCCACG	1355		
ib	995	TCCAAGCCATCAAGCTATTCAAAGAAGAAATTCTCTCCTGTGAAAGATGGGGTGAAGTTAG	1054		
iy	1356	TCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCA	1415		
ib	1055	TGAATGATCTGGCCCAACAGCTTGCCATCTCTGATGTGCATTTGTCAATGGAGAAATCCA	1114		
iy	1416	GCACCTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGGTGCCGCTCGAGGACCGAG	1475		
ib	1115	GGGCTCTGGAACAGATCAACATCCGGTGGAAACAGCTCCAGGTGTCACTGGCTGAGAGGC	1174		

QY	1476	TCAGGCAGCTGCATGAAGCCACAGGAGCTTTTGGTCCAGCATCTCAGCACTTTCTTTCCA	1535					
Db	1175	TTAAGCAACTCCAGGATGCCACCGGACTTTTGGGCTGGGTACAGCACTTCTCTCCA	1234					
QY	1536	CGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCA	1595					
Db	1235	CTTCTGTCCAAGTTCCCTGGGAAAAGAGCATTTTCTCCCAATAAAGTCCCTACTACATCA	1294					
QY	1596	ACCACGAGACTCAAAACAATTGCTGGGACCATCCCAAAAATGACAGAGCTTACCAGTCTT	1655					
Db	1295	ACCACAGGCTCAGACCACATGCTGGGACCATCTTAAGATGACTGAGTTATACCAACCC	1354					
QY	1656	TAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAC	1715					
Db	1355	TAGCCGATCTGAACAACATTAAGTTCTCAGCTTACCGCACTGCCATGAAGCTCCGCAGAG	1414					
QY	1716	TGCAGAAGGCCCTTTTGTCTGGATCTCTTGAGCCTGTGTCAGCTGTCATGTGATGCTTGGACC	1775					
Db	1415	TCCAGAAGGCCCTGCGCCTGGATCTGGTAACTTTAACTACAGCTCTGGAGATCTTCAATG	1474					
QY	1776	AGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGA	1835					
Db	1475	AGCATGACTTGACGCCAGTGAACATGTGATGGATGTGGTGGAGTCAATCACTGCTTGA	1534					
QY	1836	CCACTATTATGACCGCCTCGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCG	1895					
Db	1535	CTGCTTGTATGAACGACTGGAGGAGAAAGGSCATCTCTGTCATATGTGCCACTGTGTG	1594					
QY	1896	TGGATATGTGTCTGAACCTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCC	1955					
Db	1595	TAGACATGAGCCTCAACTGGTCTCTCAATGTTTGTATAGTGTGCGCAGTGGAAAGATGC	1654					
QY	1956	GTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGT	2015					
Db	1655	GAGCATGTCTCTTAAAGACTGGCATCGCATGCCTGTGTGGCACGGAAGTGAAGAAAC	1714					
QY	2016	ACAGATACTTTTCAAGCAACTGGCAAGTTTCAACAGGATTTTGTGACCGCGCAGGCTGG	2075					
Db	1715	TTCAGTATCTCTTCAGCCAAGTAGCCAATTTCAAGCAGCCAGTGTGATCAACGCCATCTCG	1774					
QY	2076	GCCTCCTCTGTCATGATTCTATCCAA	2101					
Db	1775	GTGCCCTGCTTCATGAAGCCATTCAA	1800					

RESULT 6
CB177816/c
LOCUS
DEFINITION
CB177816
is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
CB177816
CB177816.1 GI:28186206
EST.
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 595)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812

Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
Location/Qualifiers
1. .595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

IGIN

Query Match 24.2%; Score 509; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 7.8e-108;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAGCTATGCTACACACAGGCTGTTATGTCACCACTCTGACCCCTACCGAGCCCAT 60
|||||
509 GAGCTATGCTACACACAGGCTGTTATGTCACCACTCTGACCCCTACCGAGCCCAT 450
|||||
61 TCCTTCACAGCATTTGGAGCTCTGAAGACAAGTCATTTGGCAGTTTCATGATGGAGAG 120
|||||
449 TCCTTCACAGCATTTGGAGCTCTGAAGACAAGTCATTTGGCAGTTTCATGATGGAGAG 390
|||||
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTCTTTC 180
|||||
389 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTCTTTC 330
|||||
181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
|||||
329 TGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 270
|||||
241 CCAGTTTCATCTCATGAGGGGTACATGATGATTGACAGCCCATCAGGCGCGGTGG 300
|||||
269 CCAGTTTCATCTCATGAGGGGTACATGATGATTGACAGCCCATCAGGCGCGGTGG 210
|||||
301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
|||||
209 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 150
|||||
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAACATGGGAATGCCTCAGGGTAGC 420
|||||
149 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAACATGGGAATGCCTCAGGGTAGC 90
|||||
421 TAGCATGGAAGAAAAACAAAGCAATTACATAGAGTTTAAATGGATCTCCAGAAATCAGAACT 480
|||||
89 TAGCATGGAAGAAAAACAAAGCAATTACATAGAGTTTAAATGGATCTCCAGAAATCAGAACT 30
|||||
481 GAAAGAGTTGAATGACTGGCTAACAAAAA 509
|||||
29 GAAAGAGTTGAATGACTGGCTAACAAAAA 1

ISULT 7

653550
CUS
AGENCOURT_14553000 NIA Human H1 Embryonic Stem Cell cDNA Library

777 bp mRNA linear EST 18-JUN-2003

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

1. .777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30426742"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
(Long)"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with TRIzol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199] Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTTCTAGATCGGAGCGGCCCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lone-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Query Match 24.0%; Score 505; DB 14; Length 777;
Best Local Similarity 100.0%; Pred. No. 7.1e-107;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1597 CCACGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 1656
137 CCACGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 196
1657 ACCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAATACTCCGAGACT 1716
197 ACCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAATACTCCGAGACT 256
1717 GCAGAAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCA 1776
257 GCAGAAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCA 316
1777 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 1836
317 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 376
1837 CACTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1896
377 CACTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 436
1897 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCG 1956
437 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCG 496
1957 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
497 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 556
2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGG 2076
557 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGG 616
2077 CCTCCTTCTGCATGATTCTATCCAA 2101
617 CCTCCTTCTGCATGATTCTATCCAA 641

RESULT 8
LOCUS AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens EST 31-MAY-2003
DEFINITION cDNA clone CS0DK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK001CA09QP1.
Location/Qualifiers
source 1...824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 24.0%; Score 505; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 7.2e-107;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1597 CCACGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 1656
Db 195 CCACGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 254
QY 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAATACTCCGAGACT 1716
Db 255 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAATACTCCGAGACT 314
QY 1717 GCAGAAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCA 1776
Db 315 GCAGAAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCA 374
QY 1777 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 1836
Db 375 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 434
QY 1837 CACTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1896
Db 435 CACTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 494
QY 1897 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCG 1956
Db 495 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCG 554
QY 1957 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
Db 555 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 614
QY 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGG 2076
Db 615 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGG 674
QY 2077 CCTCCTTCTGCATGATTCTATCCAA 2101
Db 675 CCTCCTTCTGCATGATTCTATCCAA 699

RESULT 9
LOCUS CB991394 801 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT 13627932 NIH MGC 148 Homo sapiens cDNA clone IMAGE:30336570 5', mRNA sequence.
ACCESSION CB991394
VERSION CB991394.1 GI:30285818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 801)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1098)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSI1AK001ZB11QP1.
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

Query Match	24.0%;	Score 503.4;	DB 14;	Length 801;
Best Local Similarity	99.8%;	Pred. No. 1.7e-106;		
Matches 504;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
1597	CCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT	1656		
140	CCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT	199		
1657	AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT	1716		
200	AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT	259		
1717	GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA	1776		
260	GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA	319		
1777	GCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGAC	1836		
320	GCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGAC	379		
1837	CACATATTATGACCGCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGT	1896		
380	CACATATTATGACCGCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGT	439		
1897	GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCG	1956		
440	GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCG	499		
1957	TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA	2016		
500	TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA	559		
2017	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCAGGCTGGG	2076		
560	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCAGGCTGGG	619		
2077	CCTCCTTCTGCATGATTCATCCAA	2101		
620	CCTCCTTCTGCATGATTCATCCAA	644		

BX365572	1098 bp	mRNA	linear	EST 05-MAY-2003
UCUS				
INITIATION	BX365572	Homo sapiens HELA CELLS	COT 25-NORMALIZED	Homo sapiens
		cDNA clone CS0DK001YB21	5-PRIME,	mRNA sequence.

EX-100
EX-100

RESULT 13

BQ640063
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 495 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. .620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized,
unamplified): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor
[5'-pGACTAGTTCTAGATCGGAGCGGCCGCC(T)15-3']. EST analysis
was performed on the unamplified library at the NIH
Intramural Sequencing Center (NISC)."

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e-104;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1608 AAACAACCTTGCTGGACCATCCCAAATGACAGAGCTCTACCGACTCTTTAGTGACCTGA 1567
1 AAACAACCTTGCTGGACCATCCCAAATGACAGAGCTCTACCGACTCTTTAGTGACCTGA 60
1668 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGCCAATCCGAAGACTCGAAGAGGCC 1727
61 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGCCAATCCGAAGACTCGAAGAGGCC 120
1728 TTTGCTTGGATCTCTTTGAGCCTGTGACGCTGCATGTGATGCGCTTGGACCGACACACCTCA 1787
121 TTTGCTTGGATCTCTTTGAGCCTGTGACGCTGCATGTGATGCGCTTGGACCGACACACCTCA 180
1788 AGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAAATTGTTGACCACCTATTATG 1847

181 AGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAAATTGTTGACCACCTATTATG 240
1848 ACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTGC 1907
241 ACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTGC 300
1908 TGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTCTT 1967
301 TGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTCTT 360
1968 TTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTT 2027
361 TTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTT 420
2028 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGCCTCCTTCTGC 2087
421 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGCCTCCTTCTGC 480
2088 ATGATTCTATCCAA 2101
481 ATGATTCTATCCAA 494
CE960722 797 bp mRNA linear EST 29-APR-2003
AGENCOURT_13761995 NIH_MGC 147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
CE960722
CE960722.1 GI:30216839
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 797)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers
1. .797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-xhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

FEATURES
source
High quality sequence stop: 620.
Location/Qualifiers
1. .797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-xhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN
Query Match 23.5%; Score 494; DB 14; Length 797;

```
Best Local Similarity 99.8%; Pred. No. 2.6e-104;
Matches 505; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
y 1597 CCACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 1656
b 168 CCACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 227
y 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716
b 228 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 287
y 1717 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGTCATGTCATGTCCTTGACCA 1776
b 288 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGTCATGTCCTTGACCA 347
y 1777 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 1836
b 348 GCACAAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 407
y 1837 CACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTGGTCAACGTCCTCTCTGCGT 1896
b 408 CACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTGGTCAACGTCCTCTCTGCGT 467
y 1897 GGATATGTCTGTAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
b 468 GGATATGTCTGTAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 527
y 1957 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
b 528 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 587
y 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGG 2076
b 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGG 647
y 2077 -CCTCCTTCTGCATGATTTCTATCCAA 2101
b 648 CCTCCTTCTGCATGATTTCTATCCAA 673
```

```
RESULT 13
Y408546
LOCUS
DEFINITION
AY408546 Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY408546 genomic survey sequence.
AY408546 AY408546.1 GI:39764517
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2874)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2874)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1 .2874
FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2874
/gene="DRP2"
/locus_tag="HCM3260"

ORIGIN

Query Match 23.1%; Score 484.6; DB 29; Length 2874;
Best Local Similarity 58.3%; Pred. No. 5.4e-102;
Matches 784; Conservative 0; Mismatches 561; Indels 0; Gaps 0;

QY 756 ACCGCTGGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACCTTCTCTGC 815
Db 275 ACCTCCGGCTCGCCTAGAGGCTTCTCAGACCACTGGAAGCTTCAGCTCCCTCTTC 334
QY 816 AGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTG 875
Db 335 AAGAGATTATTGACTGGCTCAGCCAAAGGATGAGGAGTTGTGAGCTCAGCTGCCCCCTAC 394
QY 876 GAGGCGACTTTCAGCAGTTTCAGAAGCAGAACGATGATCATAGGGCCCTTCAAGAGGGAAT 935
Db 395 AGGGGATGTGGCCCTGGTGTCAACAGGAGAGGAGACACATGCGGCCCTTTATGGAAGAAG 454
QY 936 TGAACACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTAGGAATATTCTTGACAG 995
Db 455 TCAAGTCTCGGGGCCCTACATCTATTCTGTGTGAGTCAAGCTCAGGCCCTCTGTGCC 514
QY 996 AGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCCCAGAGAGCTGCCCTCTGAGG 1055
Db 515 AGCACCCATTGAGGAGTTAGAGGAGCCTCATTTCTGAGAGCAAAAGATACCTCCCCGAAAC 574
QY 1056 AGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGT 1115
Db 575 AGCGGATCCAGAAATCTCAGCCGCTTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACGTGT 634
QY 1116 GGGAAAAATTGAACCTGCACCTCGGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAA 1175
Db 635 GGGAGAAAGTTGACAGCCCGCTGTGTGGACCAGCACCCGTCACTTGAGCGGACTCTGGAGC 694
QY 1176 GACTCCAGGAACTTCAAGAGGCGCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGG 1235
Db 695 AGCTCTTGAGAGATTCAAGGGGCAATGGAGGAACCTAGCACTACTCTGAGCCAAAGCTGAGG 754
QY 1236 TGATCAAGGGATCCTGGCAGCCGCTGGGCGGATCTCTCATTTGACTCTCTCCAAGATCACC 1295
Db 755 GAGTCCGAGCCACTTGGGAGCCCATTTGGGATCTCTTCAATTGATTCACTCCCAAGACACA 814
QY 1296 TCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACG 1355
Db 815 TCCAGGCTATTAAAGCTGTTCAAAGAAAGAAATTCCTCCCATGAAAGATGGAGTAAAGTTGG 874
QY 1356 TCAATGACCTTGTCTCGCCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATAACCTCA 1415
Db 875 TGAATGATCTGGCCCAACCAACTGCCATTTCTGATGTGCACTTGTCAATGGAGAAATCCC 934
QY 1416 GCACCTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAG 1475
Db 935 AGGCCCTGGAACAGATCAACGTCGATGGAACAACACTACAGGCGTCAGTTAGTGAGAGGC 994
QY 1476 TCAGGCAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCA 1535
Db 995 TTAAGCAGTCCAGGATGCCCAACGGGACTTTGGGCTGGTCAAGCACTTTCTCTCCT 1054
QY 1536 CGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCA 1595
Db 1055 CCTCTGTCCAGGTTCCCTGGGAAAGAGCAATTTACCCTAAAGTTCCCTACTACATCA 1114
QY 1596 ACCAGGAGACTCAACAACACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTT 1655
Db 1115 ANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1174
QY 1656 TAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAC 1715
```

Db 1175 NNCTGATCTGAACACATTAAAGTTCTCAGCTTATCGCACTGCCATGAACCTCCGAGAG 1234
QY 1716 TGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGGATGTGATGCCCTTGAGCC 1775
Db 1235 TCCAGAAAGCCCTGCGCTTGGACCTGGTAACTTTAACACAGCCCTGGAAATCTTCAATG 1294
QY 1776 AGCACAAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCGAGATTATTAAATTGTTGA 1835
Db 1295 AGCATGATCTGCAGSCCAGTGAGCACGATGATGTGGTAGAGGTCAATTCACCTGCTGA 1354
QY 1836 CCACATATTATGACCSCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCCTCTGCG 1895
Db 1355 CTGCCTTATATGAACGTTTGGAGGAGGAAGAGGCATCCTGGTCAACGTCGCCACTCTGTG 1414
QY 1896 TGGATATGTGCTGACTGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCC 1955
Db 1415 TGGACATGAGCCTCAATGGCTCCTCAATGTTTGTATAGTGTGCGACGCGGAAGATGC 1474
QY 1956 GTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGT 2015
Db 1475 GGGCATTGTCTTTTAAAGACTGGCATTGCGATGCTTGTGTGGCAGCGAAGTGAAGGAAAAAC 1534
QY 2016 ACAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACCGAGCGAGGCTGG 2075
Db 1535 TTCAGTACCTCTTCAGCCAAAGTGGCCAACTCAGGCAGCCAGTGTGACCGGCCACCTTG 1594
QY 2076 GCCTCCTCTCTGATGATCTATCCA 2100
Db 1595 GTGTCTCTGCTTTCATGAGGCCATTCA 1619

RESULT 14
CD701871
LOCUS EST18395 human nasopharynx Homo sapiens cDNA, mRNA linear EST 25-JUN-2003
DEFINITION CD701871
ACCESSION CD701871.1 GI:32232501
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 508)
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
Location/Qualifiers
1. .508
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN
Query Match 23.0%; Score 483.8; DB 14; Length 508;
Best Local Similarity 99.6%; Pred. No. 5.7e-102;
Matches 485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1615 TTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGT 1674
Db 16 TTGCTGGGACCATCCCAAAATGACAGAGCTCTTACAGTCTTTAGCTGACCTGAATAATGT 75

QY 1675 CAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACTGCAGAGGCCCTTTTGCTT 1734
Db 76 CAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACTGCAGAGGCCCTTTTGCTT 135
QY 1735 GGATCTCTTGAGCCCTGTGAGCTGCAATGTCATGTCCTTGGACCGACACACCTCAAGCAAAA 1794
Db 136 GGATCTCTTGAGCCCTGTGAGCTGCAATGTCATGTCCTTGGACCGACACACCTCAAGCAAAA 195
QY 1795 TGACCAAGCCCATGGATATCCTGCGAGATTATTAAATTGTTGACCACTATTATGACCGCCT 1854
Db 196 TGACCAAGCCCATGGATATCCTGCGAGATTATTAAATTGTTGACCACTATTATGACCGCCT 255
QY 1855 GGACCAAGAGCACACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGCTGAACTG 1914
Db 256 GGACCAAGAGCACACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGCTGAACTG 315
QY 1915 GCTCTGATGATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAAC 1974
Db 316 GCTCTGATGATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTTATAAAC 375
QY 1975 TGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCA 2034
Db 376 TGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCA 435
QY 2035 AGTGGCAAGTTCACAGGATTTTGTGACCGCGAGGCTGGGCTCTTCTGATGATTC 2094
Db 436 AGTGGCAAGTTCACAGGATTTTGTGACCGCGAGGCTGGGCTCTTCTGATGATTC 495
QY 2095 TATCCAA 2101
Db 496 TATCCAA 502

RESULT 15
CB228986
LOCUS AGENCOURT 11499247 NICHDRh_Ov1 Macaca mulatta cDNA clone
DEFINITION CB228986
ACCESSION CB228986.1 GI:28280564
VERSION
KEYWORDS
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
cDNA Library Preparation: CLONTECH
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers
1. .728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDRh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;

Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

RIGIN

Query Match		22.8%;	Score 480;	DB 14;	Length 728;
Best Local Similarity		90.3%;	Pred. No. 4.8e-101;		
Matches 513;		Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0;
y	295	GGTTGGTAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGA	354		
b	1	GGTTGGTAATATCTACAATTGGGAAGTCAGCTGATTGGAACAGGGAATTATCAGAAGA	60		
y	355	TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG	414		
b	61	TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG	120		
y	415	GGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA	474		
b	121	GGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA	180		
y	475	GAAACTGAAAGAGTTGAATGACTGGCTAACAACAAACAGAGAAAGAACAGGAAAAATGGA	534		
b	181	GAAACTGAAAGAGTTGAATGACTGGCTGACAAAACAGAGAAAGAACAGGAAAAATGGA	240		
y	535	GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACACATAAGGT	594		
b	241	GAAAGAACCCCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACACATAAGGT	300		
y	595	GCTTCAAGAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTCACTCAGATGGTGGT	654		
b	301	GCTTCAAGAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTCACTCAGATGGTGGT	360		
y	655	GGTAGTTGATGAATCTAGTGGAGATCAGGCAACTGCTGTTTGGAGAACAACTTAAGGT	714		
b	361	GGTAGTTGATGAATCTAGTGGAGATCAGGCAACTGCTGTTTGGAGAACAACTTAAGGT	420		
y	715	ATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACA	774		
b	421	ATTGGGAGATCGATGGGCAACATCTGCAGATGGACAGAACCCGCTGGGTCTTTTACA	480		
y	775	AGACAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCT	834		
b	481	AGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTCTAGTGCATGGCT	540		
y	835	ACAGCTGAAAGATGATGAATTAAGCCGG	862		
b	541	TTCAGAAAAGAGAGATGTGGTGACAGG	568		

earch completed: April 5, 2004, 08:15:16
ob time : 3676.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

4 nucleic - nucleic search, using sw model

an on: April 4, 2004, 16:58:13 ; Search time 101.891 Seconds
(without alignments)
11443.127 Million cell updates/sec

File: US-09-845-416-12_COPY_900_3000
Effect score: 2101
Sequence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 682709 seqs, 277475446 residues

otal number of hits satisfying chosen parameters: 1365418

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB	ID	Description
1	1328.8	63.2	5952	4	US-09-687-875A-1	Sequence 1, Appli
2	1317.8	62.7	13977	4	US-09-484-970B-60	Sequence 60, Appl
3	1149.6	54.7	19307	3	US-08-836-022A-10	Sequence 10, Appl
4	1149.6	54.7	19307	3	US-09-427-048A-10	Sequence 10, Appl
5	547.4	26.1	6045	4	US-09-091-501B-7	Sequence 7, Appli
6	547.4	26.1	10320	4	US-09-091-501B-9	Sequence 9, Appli
7	517.8	24.6	3915	4	US-09-976-594-93	Sequence 93, Appl
8	79.4	3.8	200	4	US-09-091-501B-5	Sequence 5, Appli
9	78.6	3.7	200	4	US-09-091-501B-4	Sequence 4, Appli
10	78.6	3.7	200	4	US-09-091-501B-6	Sequence 6, Appli
11	76.6	3.6	7218	1	US-08-232-463-14	Sequence 14, Appl
12	46.2	2.2	505	4	US-09-621-976-15639	Sequence 15639, A
13	44.2	2.1	2574	4	US-09-668-313A-10	Sequence 10, Appl
14	44	2.1	1230025	4	US-09-198-452A-1	Sequence 1, Appli
15	43.4	2.1	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
16	42.8	2.0	1690	4	US-09-620-312D-69	Sequence 69, Appl
17	42.8	2.0	7812	3	US-09-368-590-1	Sequence 1, Appli
18	40.4	1.9	832	4	US-09-621-976-2813	Sequence 2813, Ap
19	40.4	1.9	2223	1	US-08-257-073-4	Sequence 4, Appli
20	39.2	1.9	16995	4	US-08-961-527-82	Sequence 82, Appl
21	38.6	1.8	1995	1	US-08-425-069-3	Sequence 3, Appli
22	38.6	1.8	1995	2	US-08-317-844B-3	Sequence 3, Appli
23	38.4	1.8	7672	4	US-09-220-132-24	Sequence 24, Appl
24	38.2	1.8	428	4	US-09-668-313A-3	Sequence 3, Appli
25	38.2	1.8	1131	6	5180810-3	Patent No. 5180810
26	38.2	1.8	1784	6	5180810-2	Patent No. 5180810
27	38.2	1.8	4439	4	US-09-668-313A-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-09-687-875A-1

; Sequence 1, Application US/09687875A

; Patent No. 6544786

; GENERAL INFORMATION:

; APPLICANT: Xiao, Paul

; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PEI

; FILE REFERENCE: 00792

; CURRENT APPLICATION NUMBER: US/09/687,875A

; CURRENT FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 60/158,868

; PRIOR FILING DATE: 1999-10-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 5952

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (2897)..(2898)

; OTHER INFORMATION: S4 junction site

; NAME/KEY: misc feature

; LOCATION: (3198)..(3199)

; OTHER INFORMATION: S2 junction site

US-09-687-875A-1

Query Match 63.2%; Score 1328.8; DB 4; Length 5952;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy	762	GGGTTCTTTTACAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC	821	Sequence 76, Appl
Db	3284	GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC	3343	Patent No. 5210183
Qy	822	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG	881	Sequence 22, Appl
Db	3344	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG	3403	Sequence 22, Appl
Qy	882	ACTTTCAGCAGTTTCAGAACGACGATGATCATAGGGCCTTCAAGAGGGAATTGAAA	941	Sequence 263, App
Db	3404	ACTTTCAGCAGTTTCAGAACGACGATGATCATAGGGCCTTCAAGAGGGAATTGAAA	3463	Sequence 17, Appl
Qy	942	CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	1001	Sequence 14, Appl
Db	3464	CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	3523	Sequence 23, Appl
Qy	1002	CTTTGGAAGGACTAGAGAACTCTACCAGAGCCACGAGAGCTGCTCTTGAGGAGAGAG	1061	Sequence 1, Appli

3524 CTTTGGAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCTCTGAGGAGAGAG 3583
1062 CCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGAAA 1121
3584 CCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGAAA 3643
1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 1181
3644 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 3703
1182 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAGCTCGGCCAAGCTGAGGTGATCA 1241
3704 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAGCTCGGCCAAGCTGAGGTGATCA 3763
1242 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
3764 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 3823
1302 AAGTCAAGSCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGTAGGCCACGTCAATG 1361
3824 AAGTCAAGSCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGTAGGCCACGTCAATG 3883
1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 1421
3884 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 3943
1422 TGGAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGC 1481
3944 TGGAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGC 4003
1482 AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCTCCACGTCTG 1541
4004 AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCTCCACGTCTG 4063
1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACCACG 1601
4064 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACCACG 4123
1602 AGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 1661
4124 AGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 4183
1662 ACCTGAATAATGTGATGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAGACTGCAGA 1721
4184 ACCTGAATAATGTGATGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAGACTGCAGA 4243
1722 AGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCCAGCACA 1781
4244 AGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCCAGCACA 4303
1782 ACCTCAAGCAAAATGACAGCCCATGGATATCTGCGAGATTAATAATGTTTGACCACTA 1841
4304 ACCTCAAGCAAAATGACAGCCCATGGATATCTGCGAGATTAATAATGTTTGACCACTA 4363
1842 TTTATGACCGCCTGGAGCAAGAGCAACAATAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 1901
4364 TTTATGACCGCCTGGAGCAAGAGCAACAATAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 4423
1902 TGTGCTGAACTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGTCC 1961
4424 TGTGCTGAACTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGTCC 4483
1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAACTGATACAGAT 2021
4484 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAACTGATACAGAT 4543
2022 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 2081
4544 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 4603
2082 TTCTGCATGATTTCTATCCAA 2101
4604 TTCTGCATGATTTCTATCCAA 4623

RESULT 2
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 62.7%; Score 1317.8; DB 4; Length 13977;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTTCTCTGCAGGAAC 821
Db |||||
Qy 8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAAC 8657
Db |||||
Qy 822 TTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGSCAGGCACCTATTGGAGCG 881
Db |||||
Qy 8658 TTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGSCAGGCACCTATTGGAGCG 8717
Db |||||
Qy 882 ACTTTCCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAA 941
Db |||||
Qy 8718 ACTTTCCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAA 8777
Db |||||
Qy 942 CTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGATATTTTCTGACAGAGCAGC 1001
Db |||||
Qy 8778 CTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGATATTTTCTGACAGAGCAGC 8837
Db |||||
Qy 1002 CTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAGGAGAGAG 1061
Db |||||
Qy 8838 CTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAGGAGAGAG 8897
Db |||||
Qy 1062 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 1121
Db |||||
Qy 8898 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 8957
Db |||||
Qy 1122 AATTGAACTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC 1181
Db |||||
Qy 8958 AATTGAACTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC 9017
Db |||||
Qy 1182 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCA 1241
Db |||||
Qy 9018 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCA 9077
Db |||||
Qy 1242 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
Db |||||
Qy 9078 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 9137
Db |||||
Qy 1302 AAGTCAAGGCACCTTCAGAGGAGAAATTTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATG 1361
Db |||||
Qy 9138 AAGTCAAGGCACCTTCAGAGGAGAAATTTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATG 9197
Db |||||
Qy 1362 ACCTTGCTCGCCAGCTTACCACCTTTTGGGCANTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db |||||

||||| 9198 ACCTTGCTGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCCTC 9257
||||| 1422 TGGAAGACCTGAACACACCATGGAAGCTTCTCAGAGTGGCCGTCGAGGACCGAGTCAGGC 1481
||||| 9258 TGGAAGACCTGAACACACCATGGAAGCTTCTCAGAGTGGCCGTCGAGGACCGAGTCAGGC 9317
||||| 1482 AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCAGCTTCTTCCAGCTCTG 1541
||||| 9318 AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCAGCTTCTTCCAGCTCTG 9377
||||| 1542 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 1601
||||| 9378 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 9437
||||| 1602 AGACTCAAAACAACTTGCTGGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 1661
||||| 9438 AGACTCAAAACAACTTGCTGGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 9497
||||| 1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
||||| 9498 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557
||||| 1722 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGATGTCATGTGATGCCCTTGAGCCAGCACA 1781
||||| 9558 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGATGTCATGTGATGCCCTTGAGCCAGCACA 9617
||||| 1782 ACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
||||| 9618 ACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 9677
||||| 1842 TTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATA 1901
||||| 9678 TTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATA 9737
||||| 1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 1961
||||| 9738 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 9797
||||| 1962 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAAGTACAGAT 2021
||||| 9798 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAAGTACAGAT 9857
||||| 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCT-GGGCCTC 2080
||||| 9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGGCCTC 9917
||||| 2081 CTTCTGCATGATTCTATCCAA 2101
||||| 9918 CTTCTGCATGATTCTATCCAA 9938

3-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVFN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-836-022A-10
Query Match 54.7%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 762 GGGTTCTTTTACAAGACAGTTCTTGACCAGTGGAGCGTCTGCACCTTTTCTCTGCAGGAAC 821
DB 6096 GGTCCCATTTGGAAGCAAGTTCTGACCAGTGGAGCGTTTGCATCTTTCTCTTCAGGAAC 6037
QY 822 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
DB 6036 TTCTTGTTTGGCTACAGCTGAAGATGATGAATGAAGCGGTGAGGACCCCATCGTGGTG 5977
QY 882 ACTTTCCAGCAGTTTCAGAAGCAGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAA 941
DB 5976 ATTTCCAGCAGTTTCAGAAGCAGATGATATACATAGGGCCCTTCAAGAGGGGAATTGAAA 5917
QY 942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
DB 5916 CTAAAGAACCTGTAATCATGAGTACTCTGAGACTGTGAGAAATATTTCTGACAGAGCAGC 5857
QY 1002 CTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGTGCCTCCTTGAGGAGAGAG 1061
DB 5856 CTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAACTGCCTCCTTGAGAGAGAG 5797
QY 1062 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
DB 5796 CTCAGAAATGTCACTCGGCTCCTACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGACA 5737
QY 1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC 1181
DB 5736 AATTGAACCTGCCTCAGCTGATTGGCAGAGAAAAATAGATGAAGCTCTTGAAAGACTCC 5677
QY 1182 AGGAACTTCAAGAGGCCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGTGATCA 1241
DB 5676 AGGAACTTCAAGAGGAGCTGCCGATGAACCTGAGCTTGCCTCCTGAGGAGGAGTCA 5617
QY 1242 AGGATCCTGCGCAGCCCGTGGGCGATCTCCTCATTTGAGTCTCTCTCAAGATCACCTCGAGA 1301
DB 5616 AGGATCCTGCGCAGCCAGTGGGGATCTCCTCATTTGAGTCTCTCTGCAAGATCACCTTGAA 5557
QY 1302 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGCTGAGCCACCTCAATG 1361
DB 5556 AAGTCAAGGCACTTCGGGGAGAAATTCACCTCTTAAAGAGAAATGTCAATCGTGTCAATG 5497
QY 1362 ACCTTGCTCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
DB 5496 ACCTTGACATCAGCTGACCACACTGGGGCAATTCAGCTCTCACCTTATAACCTCAGCACTT 5437

1422 TGGAGACCTGAACACCAGATGGAAGCTTCTGAGTGGCCGTCGAGACCGAGTCAAGC 1481
5436 TGGAGATCTGAATACCAGATGGAGGCTTCTACAGTGGCTGTGGAGACCGTGTAGAC 5377
1482 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTG 1541
5376 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACTTCAG 5317
1542 TCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCACG 1601
5316 TTCAGGTCCTGGGAGAGAGCCATCTCACCACAAACAAAGTCCCTACTATATCAACCACG 5257
1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 1661
5256 AGACCCAAACCACTTGTGGGACCAACCCAAATGACAGAGCTCTACCACTTTTAGCTG 5197
1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
5196 ACCTGAATAATGTCAGTTCTCCGGGTATAGGACTGCCATGAAGCTCAGAAGGCTCCAGA 5137
1722 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACA 1781
5136 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACA 5077
1782 ACCTCAAGCAAAATGACCCAGCCATGGATATCCTGACAGATTATTAATTGTTGACCACCTA 1841
5076 ACCTCAAGCAAAATGACCCAGCCATGGATATCCTGACAGATAATTAATCTGTTGACTACAA 5017
1842 TTTATGACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTGCGTGGATA 1901
5016 TTTATGATCGCTGGAGCAAGAGCAACAAATCTGGTCAATGTCTCTCTCTGTTGGATA 4957
1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
4956 TGTGCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4897
1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAGAGCAAGTACAGAT 2021
4896 TGTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCAACATTTGGAAGCAAGTACAGAT 4837
2022 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTGTGACCGCAGCGTGGGCTTCC 2081
4836 ACCTTTTCAAGCAAGTGGCAAGTTTCAACTGGCTTTTGTGACCGCAGCGTAGGCTGGTCTTC 4777
2082 TTCTGCATGATTTCTATCCAA 2101
4776 TTCTGCATGATTTCTATTCAA 4757

ESULT 5
S-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 26.1%; Score 547.4; DB 4; Length 6045;
Best Local Similarity 61.4%; Pred. No. 5.3e-160;
Matches 900; Conservative 0; Mismatches 556; Indels 9; Gaps 1;
QY 645 ACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAGAAC 704
Db 3270 ACAGGCAGAGATGGTAAAGCTTTGGGAAATTTCTGAAGAGCTACTATGCTTCAACATC 3329
QY 705 AACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGG 764
Db 3330 GACTGGATGATATGAACCAAGATGAATGACTTAAAGCAAAATCTGCTAGCATCAGGG 3389
QY 765 TTCTTTTACAGACAGTTCTTGACCAGTGGAAAGCGTCTGACCTTTCTCTGCAGGAACCTTC 824
Db 3390 CCCATTTGGAGGCCAGCGCTGAGAAGTGGAAACAGGTTGCTGATGTCCTTAGAAGAACTGA 3449
QY 825 TGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGGACT 884
Db 3450 TCAATGGCTGAATATGAAGATGAAGAGCTTAAGAAACAAATGCCTATTGGAGGAGATG 3509
QY 885 TTCCAGCAGTTTCAGAAAGCAGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAACTA 944
Db 3510 TTCCAGCCTTACAGTCCAGTATGACCATTTGTAAGGCCCTGAGACGGGAGTTAAAGGAGA 3569
QY 945 AAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTT 1004
Db 3570 AAGAATATTCTGTCTCTGAATGCTGTGACCCAGGCCGAGTTTCTTGGCTGATCAGCCAA 3629
QY 1005 T-----GGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGG 1055
Db 3630 TTGAGGCCCTTGAAGAGCCAGAGAAACCTTACAATCAAAAACAGAAATTAACCTCCTGAGG 3689
QY 1056 AGAGAGCCCAAGATGTCACTCGGCTTTTACGAAAGCAGGCTGAGGAGGTCAATACTGAGT 1115
Db 3690 AGAGAGCCCAAGATGTCCAAAGCCATGCGCAACACAGTCTTCTGAAAGTCAAGAAAT 3749
QY 1116 GGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAA 1175
Db 3750 GGGAAAGTCTAATGCTGTAACTAGCAATTTGGCAAGCAAGTGGACAGGCATTGGAGA 3809
QY 1176 GACTCCAGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGG 1235
Db 3810 AACTCAGAGACCTGCAGGAGCTATGGATGACCTGGACGCTGACATGAAGGAGGCAGAGT 3869
QY 1236 TGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACC 1295
Db 3870 CCGTGGGAATGGCTGGAAAGCCCGTGGGAGACTTACTCATTTGACTCGCTGAGGATCACA 3929
QY 1296 TCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCCCTCTGAAAGAGAACCGTGAGCCACG 1355
Db 3930 TTGAAAAAATCATGGCAFTTAGAGAAGAAATTTGCACCAATCAACTTTAAAGTTAAACGG 3989
QY 1356 TCAATGACCTTGTCTGCCAGCTTACCATTGCGCATTCAGCTCTCACCGTATAACCTCA 1415
Db 3990 TGAATGATTTATCCAGTCAGCTGTCTCCACTTGACCTGCATCCCTCTCTAAAGATGTCTC 4049
QY 1416 GCACCTCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGCTCAGGACCGAG 1475
Db 4050 GCCAGCTAGATGACCTTAATATGCGATGGAACCTTTACAGGTTTCTGTGATGATCGCC 4109

Y	1536	CGTCTGCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCA	1595
b	8445	CGTCAGTCCAGTGCCTGGCAAGATCCATTTTCACATAATAAAGTGCCCTATTACATCA	8504
Y	1596	ACCACGAGACTCAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGICTT	1655
b	8505	ACCATCAAAACACAGACCACCTGTTGGACCATCTTAAATGACCGAATCTTTCAATCCC	8564
Y	1656	TAGCTGACCTGAATAATGTACAGATTCTCAGCTTATAGGACTGCCATGAATCCCGAAGAC	1715
b	8565	TTGCTGACCTGAATAATGTACGTTTTTCTGCTTACCGTACAGCAATCAAAATCCGAAGAC	8624
Y	1716	TGCAGAAGGCCCTTTTGCTTGGATCTCTTGAGCCCTGTTCAGCTGCATGTGATGCCCTTGGACC	1775
b	8625	TACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACAAACAATGAATTTTCAAAC	8684
Y	1776	AGCACAACCTCAAGCAAAATGACCAGCCCCATGGATATCCTGTCAGATTATTAATGTTTGA	1835
b	8685	AGCAAAAGTTGAACCAAAATGACCAGCTCCTCAGTGTTCAGATGTTCATCAACTGTCTGA	8744
Y	1836	CCACTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCG	1895
b	8745	CAACAACCTTATGATGGACTTGAGCAAAATGCATAAGGACCTGGTCAACGTTCCACTCTGTG	8804
Y	1896	TGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCC	1955
b	8805	TTGATATGTGTCTCAATTGGTTGCTCAATGTCTATGACACGGGTGCACTGGAAAAATTA	8864
Y	1956	GTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGCAAGT	2015
b	8865	GAGTGCAGAGTCTGAAGATTGGATTAAATGTCTCTCTCCAAAGGTCTCTTGGAAAGAAAAT	8924
Y	2016	ACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGG	2075
b	8925	ACAGATATCTCTTTAAGGAAGTTGCGGGCCGACAGAAAATGTGTGACAGAGGCAGCTGG	8984
Y	2076	GCCTCCTTCTGCATGATTCTATCCA	2100
b	8985	GCCTGTACTTTCATGATGCCATCCA	9009
RESULT 7			
IS-09-976-594-93			
Sequence 93, Application US/09976594			
Patent No. 6673549			
GENERAL INFORMATION:			
APPLICANT: Furness, Michael			
APPLICANT: Buchbinder, Jenny			
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS			
FILE REFERENCE: PA-0041 US			
CURRENT APPLICATION NUMBER: US/09/976,594			
CURRENT FILING DATE: 2001-10-12			
PRIOR APPLICATION NUMBER: 60/240,409			
PRIOR FILING DATE: 2000-10-12			
NUMBER OF SEQ ID NOS: 1143			
SOFTWARE: PERL Program			
SEQ ID NO 93			
LENGTH: 3915			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: misc feature			
OTHER INFORMATION: Incyte ID No. 6673549 290344.1.			
IS-09-976-594-93			
Query Match			
Best Local Similarity 24.6%; Score 517.8; DB 4; Length 3915;			
Matches 846; Conservative 0; Mismatches 527; Indels 2; Gaps 1;			
Y	726	GATGGCAACATCTGTAGATGGACAGAGACCGTGGTTCTTTTACAAGACAGTTCTG	785
b	555	GTTGGAATGAATAAAAAAAGTCTCACAACTCCGGCTCGCCTAGAGGCCTTCTCAG	614

QY	786	ACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAG	845
Db	615	ACCAAGTGAAGCGTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCTCAGCCAAAAGG	674
QY	846	ATGATGAATTAAGCCGGCAGSCACCTATTGGAGGCGACTTTCCAGCAGTTCCAGAAGCAGA	905
Db	675	ATGAGGAGTTGTGAGCTCAGCTCAGCTGCCCCCTACAGGGGGATGTGGCCCTGGTGCACAGGAGA	734
QY	906	ACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTTAAAGAACCTGTATATCATGAGTA	965
Db	735	AGGAGACACATGCGGCCCTTTATGGAAGAAGTCAAGTCTCGGGGCCCTACATCTATTCTG	794
QY	966	CTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTTAGAGAAACTCT	1025
Db	795	TGCTGGAGTCAGCTCAGGCCCTTCTGTCCCAGCACCCATTTGAGGAGTTAGAGAGCCTC	854
QY	1026	ACCAGGAGCCACAGAGAGCTGCCCTCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTAC	1085
Db	855	ATTCTGAGAGCAAAAGATACCTCCCCGAAACAGCGGATCCAGAAATCTCAGCCCGCTTGTAT	914
QY	1086	GAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACT	1145
Db	915	GGAAGCAGCGACGGTGGCCAGTGAACCTGTGGGAGAACTTGACAGCCCGCTGTGTGGACC	974
QY	1146	GGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACTTCAAGAGGCCACCGGATG	1205
Db	975	AGCACCGTCAATTTGAGCGGACTCTGGAGCAGCTCTTGGAGATTTCAGGGG--CATGGAGG	1032
QY	1206	AGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGG	1265
Db	1033	AACTAAGCACTACTCTGAGCCCAAGCTGAGGGAGTCCGAGCCACTTGGGAGCCCATTTGGGG	1092
QY	1266	ATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAA	1325
Db	1093	ATCTCTTCATTTGATTCACTCCAGAGCACATCCAGGCTATTAAAGCTGTTCAAAGAAGAAAT	1152
QY	1326	TTGCGCCTCTGAAAGAGAACGTCAGGCCACGTCATCAATGACCTTGCTCGCCAGCTTACCAC	1385
Db	1153	TCTCCCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCACTTGGCCATTT	1212
QY	1386	TGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGA	1445
Db	1213	CTGATGTGCACCTTGTCATATGGAGAAATCCCAGGCCCTGGAACAGATCAACGTCGGATGGA	1272
QY	1446	AGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACT	1505
Db	1273	AACAACCTACAGCGCTCAGTTGATGAGAGGCTTAAGCAGCTCCAGGATGCCACCGGGACT	1332
QY	1506	TTGGTCCAGCATCTCAGCACITTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCA	1565
Db	1333	TTGGGCTGGGTACAGCACITTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1392
QY	1566	TCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACC	1625
Db	1393	TTTCACCCCAATAAAGTTCCTCTACTACATCAACCACCGGCTCAGACCACATGCTGGGACC	1452
QY	1626	ATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAAATAATGTCAATTTCTCAG	1685
Db	1453	ATCCCAAGATGACAGAGTTATACCAACCCCTAGCTGATCTGAAACAACATTAAGTTCTCAG	1512
QY	1686	CTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGA	1745
Db	1513	CTTATCGCACTGCCATGAACTCCGCAGAGTCCAGAAAGCCCTGCGCTTGGACCTGGTAA	1572
QY	1746	GCCTGTACGCTGCATGTGATGCTTGGACCGACCAACCTCAAGCAAAATGACCAGCCCCA	1805
Db	1573	CTTTAACCAACAGCCCTGGAAATCTTCAATGAGCATGATCTGCAGGCCAGTGAGCACGTGA	1632
QY	1806	TGGATATCTGCAGATTATTAAATGTTTGGACCACTATTATATGACCGCCTGGAGCAAGAGC	1865
Db	1633	TGGATGTGTAGAGGTCAATCACTGCCTGACTGCCTTATATGAACGTTTGGAGGAGGAAA	1692
QY	1866	ACAACAATTTGGTCAACGTCCTCTCTCGTGGATATGTGTCTGAACTGGCTGCTGAATG	1925

Db 1693 GAGGCATCTCGGTCAACGTGCCACTCTGTGTGGACATGAGCCTCAATTGGCTCCTCAATG 1752
2Y 1926 TTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGCTCTTTTAAAACTGGCATCATTT 1985
Db 1753 TTTTGTAGTGGTTCGCAGCGGAAGATCGGGCATTGCTCTTTTAAGACTGGCATTCAT 1812
2Y 1986 CCCTGTGTAAAGCACATTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTT 2045
Db 1813 GCTTGTGTGCACGGAAGTGAAGGAAAAAATTCAGTACCTCTTCAGCCAAAGTGGCCAACT 1872
2Y 2046 CAACAGGATTTGTGACCAGCGAGGCTGGGCCTCCTCTGTCATGATTTCTATCCA 2100
Db 1873 CAGGCAGCCAGTGTGACCAGCGCCACCTTGGTGTCTCCTGCTTCATGAGGCCAATTCA 1927

RESULT 8

JS-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
JS-09-091-501B-5

Query Match 3.8%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 9.4e-15;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
2Y 565 CCTAAACGCCAAGTACAACAATAAGGTGCTTCAAGAAGATCTAGAACAGAACAAGT 624
Db 16 CTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAGTGACCTCGAAGCTGAGCAGGT 75
2Y 625 CAGGGTCAATCTCTCACTACATGGTGGTGTGATGATGAATCTAGTGGAGATCAGCG 694
Db 76 GAAGGTGAATTCCTTAACTCATATGCTGGTGATTGTGGATGAAAACAGTGGGAGAGCGC 135
2Y 685 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG 744
Db 136 CACAGCTGTTTGGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
2Y 745 ATGGA 749
Db 196 CTGGA 200

RESULT 9

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.7e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 553 TGATCTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCAAGAAGATCTAGA 612
Db 4 TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACAATAAAAGTTTGCAAAATGACCTTGA 63
QY 613 ACAAGAACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG 672
Db 64 AGCTGAACAGGTGAAGTAAATTCCTTAACCTCACATGGTGGTGGTAGTTGATGAAGAACAG 123
QY 673 TGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
Db 124 TGGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAAACTGGGTGAGCGGTGGAC 183
QY 733 AAACATCTGTAGATGGA 749
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 10

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.7e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 553 TGATCTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCAAGAAGATCTAGA 612
Db 4 TGATGTGAAATCTCTACAAAGCTGCTAGAAGAACATAAAAGTTTGCAAAAGTGATCTTGA 63

613 ACAAGAACAAAGTCAGGTCATCTCTCACTCACATGGTGGTGGTAGTGAATCTAG 672
64 GGCTGAACAGGTGAAAGTAAATTCACAACTCACTGATGGTGGTCACTGTTGATGAAACAG 123
673 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
124 TGGTGAGAGCGCTACAGCTATCTCTAGAGACCAGTTACAGAAACTTGGTGAGCGCTGGAC 183
733 AAACATCTGTAGATGGA 749
184 AGCAGTATGCCGTTGGA 200

RESULT 11
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 3.6%; Score 76.6; DB 1; Length 7218;
Best Local Similarity 6.7%; Pred. No. 8.7e-13;
Matches 28; Conservative 236; Mismatches 155; Indels 0; Gaps 0;
213 TTCTAATGATGTGGAAGTGGTGAAGACCAAGTTTCATACTCATGAGGGGTACATGATGG 272
1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAATTGGTACRRRRRRRRRRRRRRRR 1415
273 ATTTGACAGCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAACTGATTG 332

1414 RR 1355
333 GAACAGGAAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAA 392
1354 RR 1295
393 ATTCAAGATGGGAATGCCTCAGGCTAGCTAGCATGGGAAAAACAAAGCAATTTCATAGAG 452
1294 RR 1235
453 TTTTAATGGATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG 512
1234 RR 1175
513 AAGAAAGAACAAAGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTAAAAAC 572
1174 RR 1115
573 GCCAAGTACAAACATAGGTGCTTCAAGAGATCTAGAACAAAGCAAGTCAGGCTC 631
1114 RR 1056

RESULT 12
US-09-621-976-15639/c
Sequence 15639, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 15639
LENGTH: 505
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match 2.2%; Score 46.2; DB 4; Length 505;
Best Local Similarity 17.4%; Pred. No. 0.00043;
Matches 54; Conservative 131; Mismatches 124; Indels 2; Gaps 1;
319 AAGTAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGAACTGAAGTACAAGAGCA 378
445 AATYACRWAWAGAGAAWKRWKMGSMRACARAGTTTMAWGACWAMARRGWRRCARG--T 388
379 GATGAATCTCCTAAATTCAAGATGGGATGCCTCAGGGTAGCTAGCATGGAAAAACAAAG 438
387 GSWGSKGGYRMWGRGMAAAKMRMAAAGGSGYCGMTSYTSGSKMTGRKGSMTKRKMMTY 328
439 CAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAACTGAAAGAGTTGAATGACTG 498
327 SCMMWTSYKCTKTGKKYTGWKSCKTRWTCTSWRKYMMWMSGCGWARSMSKSWARSWSYMMAC 268
499 GCTAACAAAAACAGAAGAAAGAACAGGAAATGGAGGAGAGCCTCTTGGACCTGATCT 558
267 WCMMSASAYRARRSNYGARRSVMRAGAGWWRARRGKRARGKSSMWSKMSMRMSAGKA 208
559 TGAAGACCTAAACCCCAAGTACAAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGA 618
207 RMCRRWMWSCRRMSYSYCMGSKMSCRCGTCAKMWRYARYAKRYASSMGKYMGCRCWCYAKC 148
619 ACAAGTCAGGG 629
147 ARMYGYRSRS 137

RESULT 13

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
SEQUENCE CHARACTERISTICS:
LENGTH: 1179 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1179
SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
JS-09-107-532A-1186

Query Match	2.1%;	Score 43.4;	DB 4;	Length 1179;
Best Local Similarity	50.7%;	Pred. No. 0.0058;		
Matches 104;	Conservative 0;	Mismatches 101;	Indels 0;	Gaps 0;

2y	471	ATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGGAAAA	530
db	602	AGCAAGATATAAAGATTGTGATCGATCAGACAAAGAAAAATGGAGATACGATCGGAGAA	661
2y	531	TGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACACATA	590
db	662	TTGTAGAAGTGCTCGTTGGAGGCGTTCAGCTGGATTAGGAAGCTACGTACAATGGGACA	721
2y	591	AGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGG	650
db	722	CGAAGCTAGATGCCAAAATCGCACAAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCGTAG	781
2y	651	TGGTGGTAGTTGATGAATCTAGTGG	675
db	782	AATTGGGGTCGGATTCACTTCTGG	806

Search completed: April 5, 2004, 08:25:00
Job time : 105.891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 nucleic - nucleic search, using sw model

run on: April 5, 2004, 04:53:30 ; Search time 509.455 Seconds
(without alignments)
15454.047 Million cell updates/sec

file: US-09-845-416-12_COPY_900_3000
effect score: 2101
sequence: 1 gagtatgcctacacacagg.....ttctgcatgattctatccaa 2101

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 2470430 seqs, 1873665578 residues

total number of hits satisfying chosen parameters: 4940860

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: *
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match %	Length	DB ID	Description
1	2101	100.0	3510	10	US-09-845-416-12 Sequence 12, Appl
2	2101	100.0	4476	10	US-09-845-416-31 Sequence 31, Appl
3	1760.4	83.8	5339	15	US-10-149-736-40 Sequence 40, Appl
4	1750.8	83.3	3531	10	US-09-845-416-10 Sequence 10, Appl
5	1750.8	83.3	4498	10	US-09-845-416-30 Sequence 30, Appl
6	1743	83.0	3858	10	US-09-845-416-9 Sequence 9, Appli
7	1743	83.0	4825	10	US-09-845-416-29 Sequence 29, Appl
8	1743	83.0	4848	10	US-09-845-416-35 Sequence 35, Appl
9	1743	83.0	5060	10	US-09-845-416-36 Sequence 36, Appl
10	1641	78.1	4414	10	US-09-845-416-32 Sequence 32, Appl
11	1629	77.5	3446	10	US-09-845-416-14 Sequence 14, Appl
12	1629	77.5	5462	15	US-10-149-736-41 Sequence 41, Appl
13	1611.6	76.7	5417	15	US-10-149-736-39 Sequence 39, Appl
14	1602	76.2	3999	10	US-09-845-416-6 Sequence 6, Appli
15	1602	76.2	4966	10	US-09-845-416-28 Sequence 28, Appl

16	1602	76.2	4990	10	US-09-845-416-34 Sequence 34, Appl
17	1419	67.5	4182	10	US-09-845-416-2 Sequence 2, Appli
18	1419	67.5	5149	10	US-09-845-416-27 Sequence 27, Appl
19	1328.8	63.2	2169	10	US-09-845-416-4 Sequence 4, Appli
20	1328.8	63.2	8689	15	US-10-149-736-42 Sequence 42, Appl
21	1328.8	63.2	11058	10	US-09-845-416-1 Sequence 1, Appli
22	1328.8	63.2	11443	15	US-10-149-736-44 Sequence 44, Appl
23	1328.8	63.2	12057	15	US-10-149-736-47 Sequence 47, Appl
24	1328.8	63.2	13957	9	US-09-782-378A-22 Sequence 22, Appl
25	1328.8	63.2	13957	9	US-09-880-107-2284 Sequence 2284, Ap
26	1328.8	63.2	13957	15	US-10-149-736-1 Sequence 1, Appli
27	1328.8	63.2	14069	12	US-10-342-887-434 Sequence 434, App
28	1328.8	63.2	14082	12	US-10-342-887-981 Sequence 981, App
29	1328.8	63.2	14082	15	US-10-341-434-108 Sequence 108, App
30	1324	63.0	1821	10	US-09-845-416-13 Sequence 13, Appl
31	1149.6	54.7	13815	15	US-10-149-736-2 Sequence 2, Appli
32	937	44.6	1434	10	US-09-845-416-15 Sequence 15, Appl
33	785.4	37.4	1991	10	US-09-845-416-7 Sequence 7, Appli
34	777	37.0	1667	10	US-09-845-416-3 Sequence 3, Appli
35	554.2	26.4	11096	15	US-10-149-736-4 Sequence 4, Appli
36	547.4	26.1	10302	9	US-09-782-378A-23 Sequence 23, Appl
37	547.4	26.1	10302	15	US-10-149-736-3 Sequence 3, Appli
38	538.6	25.6	16531	14	US-10-101-510-667 Sequence 667, App
39	527.4	25.1	5106	12	US-10-220-120-157 Sequence 157, App
40	450	21.4	1340	10	US-09-845-416-11 Sequence 11, Appl
41	393	18.7	887	15	US-10-149-736-35 Sequence 35, Appl
42	387	18.4	387	15	US-10-149-736-32 Sequence 32, Appl
43	331	15.8	333	15	US-10-149-736-9 Sequence 9, Appli
44	327	15.6	327	15	US-10-149-736-8 Sequence 8, Appli
45	324	15.4	324	15	US-10-149-736-33 Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match	100.0%;	Score 2101;	DB 10;	Length 3510;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GAGCTATGCCCTACACACAGGCTGCTTATGTGTCACCCACCTCTGACCCCTACACGGAGCCCAATT	60	
Db	900	GAGCTATGCCCTACACACAGGCTGCTTATGTGTCACCCACCTCTGACCCCTACACGGAGCCCAATT	959	
QY	61	TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG	120	
Db	960	TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG	1019	
QY	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	180	
Db	1020	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	1079	
QY	181	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	240	

1080 TGCTGAGGACACATTGCAAGGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAGTTTCATACACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGTTGG 300
1140 CCAGTTTCATACACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGTTGG 1199
301 TAATAATTCTACAATTGGGAAGTAAGCTGATTGGACAGGAAAAATTATCAGAAGATGAAGA 360
1200 TAATAATTCTACAATTGGGAAGTAAGCTGATTGGACAGGAAAAATTATCAGAAGATGAAGA 1259
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1379
481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 540
1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 1439
541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACACATAAAGGTGCTTCA 600
1440 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACACATAAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAAGAACAACTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 660
1500 AGAAGATCTAGAACAAGAACAACTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 1559
661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACTTAAGGTATTGGG 720
1560 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACTTAAGGTATTGGG 1619
721 AGATCGATGGGCACAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAG 780
1620 AGATCGATGGGCACAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAG 1679
781 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCT 840
1680 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCT 1739
841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGCACTTTCCAGCAGTTCAGAA 900
1740 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGCACTTTCCAGCAGTTCAGAA 1799
901 GCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCAT 960
1800 GCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCAT 1859
961 GAGTACTCTTGAGACTGTACGAATATTCTGACAGACAGCCCTTTGGAAGGACTAGAGAA 1020
1860 GAGTACTCTTGAGACTGTACGAATATTCTGACAGACAGCCCTTTGGAAGGACTAGAGAA 1919
1021 ACTCTACAGAGAGCCCGCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCT 1080
1920 ACTCTACAGAGAGCCCGCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCT 1979
1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGC 1140
1980 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGC 2039
1141 TGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCCAGGAACTTCAAGAGGCCAC 1200
2040 TGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCCAGGAACTTCAAGAGGCCAC 2099
1201 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGT 1260
2100 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGT 2159
1261 GGGCGATCTCTCATTTGACTCTCTCCAAGATCACTCCAGAAAGTCAAGGCACCTTCGAGG 1320

2160 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG 2219
1321 AGAAATTGGCGCCTCTGAAAAGAGAAACGTGAGCCACGTCATCAATGACCTTGTCTCGCCAGCTTAC 1380
2220 AGAAATTGGCGCCTCTGAAAAGAGAAACGTGAGCCACGTCATCAATGACCTTGTCTCGCCAGCTTAC 2279
1381 CACTTTGGSCATTTCAGCTCTCACCGTATAAACCTCAGACACTCTGGAAGACCTTGAACACCAG 1440
2280 CACTTTGGSCATTTCAGCTCTCACCGTATAAACCTCAGACACTCTGGAAGACCTTGAACACCAG 2339
1441 ATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAG 1500
2340 ATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAG 2399
1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1560
2400 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 2459
1561 AGCCATCTCGCCAAAACAAAGTGGCCTACTATATCAACACGAGACTCAAACAACCTTGCTG 1620
2460 AGCCATCTCGCCAAAACAAAGTGGCCTACTATATCAACACGAGACTCAAACAACCTTGCTG 2519
1621 GGACCATCCCAAAATGACAGAGTCTTACAGTCTTTTAGCTGACCTGAATAATGTCAAGATT 1680
2520 GGACCATCCCAAAATGACAGAGTCTTACAGTCTTTTAGCTGACCTGAATAATGTCAAGATT 2579
1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTTGAGAGAGCTGAGAAAGGCCCTTTGCTGGATCT 1740
2580 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTTGAGAGAGCTGAGAAAGGCCCTTTGCTGGATCT 2639
1741 CTTGAGCCTGTGAGCTGCATGTGATGCCCTTGAGCCAGCACAAACCTCAAGCAAAATGACCA 1800
2640 CTTGAGCCTGTGAGCTGCATGTGATGCCCTTGAGCCAGCACAAACCTCAAGCAAAATGACCA 2699
1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 1860
2700 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 2759
1861 AGAGCACAAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 1920
2760 AGAGCACAAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 2819
1921 GAATGTTTATGATACGGGACGGAACAGGGAGGATCCGTCCTCTCTTTTAAAACTGGCAT 1980
2820 GAATGTTTATGATACGGGACGGAACAGGGAGGATCCGTCCTCTCTTTTAAAACTGGCAT 2879
1981 CATTTCCTCTGTATAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
2880 CATTTCCTCTGTATAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2939
2041 AAGTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCTCCTTCTGTCATGATTCATATCCA 2100
2940 AAGTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCTCCTTCTGTCATGATTCATATCCA 2999
2101 A 2101
3000 A 3000

RESULT 2

US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 31

LENGTH: 4476

TYPE: DNA

ORGANISM: Homo sapiens

3-09-845-416-31

Query Match 100.0%; Score 2101; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
1656 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1715
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 120
1716 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 1775
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
1776 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1835
181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1836 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1895
241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCCGGGTTGG 300
1896 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCCGGGTTGG 1955
301 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAGATGAAGA 360
1956 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAGATGAAGA 2015
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGSCCTCAGGGTAGC 420
2016 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGSCCTCAGGGTAGC 2075
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
2076 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2135
481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
2136 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 2195
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAAATAGGTGCTTCA 600
2196 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAAATAGGTGCTTCA 2255
601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
2256 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 2315
661 TGATGAATCTAGTGGAGATCACGCCAATGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 720
2316 TGATGAATCTAGTGGAGATCACGCCAATGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 2375
721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAG 780
2376 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAG 2435
781 TTCTGACCAGTGGAGCGTCTGCACCTTCTCTGCAGGAATCTTGGTGTGGCTACAGCT 840
2436 TTCTGACCAGTGGAGCGTCTGCACCTTCTCTGCAGGAATCTTGGTGTGGCTACAGCT 2495
841 GAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 900
2496 GAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 2555
901 GCAGAACGATGTACATAGGGCTTCAAGAGCGGAATTGAAAACTAAAGAACCTGTAATCAT 960

Db 2556 GCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCAT 2615
QY 961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020
Db 2616 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 2675
QY 1021 ACTCTACCAGGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1080
Db 2676 ACTCTACCAGGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 2735
QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCG 1140
Db 2736 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCG 2795
QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1200
Db 2796 TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2855
QY 1201 GGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGT 1260
Db 2856 GGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGT 2915
QY 1261 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
Db 2916 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2975
QY 1321 AGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTAC 1380
Db 2976 AGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTAC 3035
QY 1381 CACTTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAG 1440
Db 3036 CACTTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAG 3095
QY 1441 ATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 1500
Db 3096 ATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 3155
QY 1501 GGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1560
Db 3156 GGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 3215
QY 1561 AGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGCTG 1620
Db 3216 AGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGCTG 3275
QY 1621 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGT 1680
Db 3276 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGT 3335
QY 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAGGCCCTTTGCTTGGATCT 1740
Db 3336 CTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAGGCCCTTTGCTTGGATCT 3395
QY 1741 CTTGAGCCTGTGAGTGCATGTGATGCCCTTGACCCAGCAACCTCAAGCAAAAATGACCA 1800
Db 3396 CTTGAGCCTGTGAGTGCATGTGATGCCCTTGACCCAGCAACCTCAAGCAAAAATGACCA 3455
QY 1801 GCCCATGGATATCCTGCAGATTATTAATTTGACCCACTATTTATGACCGCCTGGAGCA 1860
Db 3456 GCCCATGGATATCCTGCAGATTATTAATTTGACCCACTATTTATGACCGCCTGGAGCA 3515
QY 1861 AGAGCACAACTTTGGTCAACGTCCCTCTCTCGTGGGATATGTCTGAACTGGCTGCT 1920
Db 3516 AGAGCACAACTTTGGTCAACGTCCCTCTCTCGTGGGATATGTCTGAACTGGCTGCT 3575
QY 1921 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTTCTTTAAAACTGGCAT 1980
Db 3576 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTTCTTTAAAACTGGCAT 3635
QY 1981 CATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db 3636 CATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 3695

1606 TCAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 1665
2719 TCAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 2778
1666 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAGGC 1725
2779 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAGGC 2838
1726 CCTTTGCTTGGATCTCTTGAGCTGTTCAGCTGCATGTGATGCCCTTGGACCAACCAACT 1785
2839 CCTTTGCTTGGATCTCTTGAGCTGTTCAGCTGCATGTGATGCCCTTGGACCAACCAACT 2898
1786 CAAGCAAAATGACCAAGCCCATGATATCCTGCAGATTATTAATTGTTTGACCACCTATTTA 1845
2899 CAAGCAAAATGACCAAGCCCATGATATCCTGCAGATTATTAATTGTTTGACCACCTATTTA 2958
1846 TGACCCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTG 1905
2959 TGACCCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTG 3018
1906 TCTGAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGGAGGATCCGTGTCCTGTC 1965
3019 TCTGAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGGAGGATCCGTGTCCTGTC 3078
1966 TTTTAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGCAAGTACAGATACCT 2025
3079 TTTTAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGCAAGTACAGATACCT 3138
2026 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCTCTTCT 2085
3139 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCTCTTCT 3198
2086 GCATGATTCATCCAA 2101
3199 GCATGATTCATCCAA 3214

3-09-845-416-10
Sequence 10, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3531
TYPE: DNA
ORGANISM: Homo sapiens
3-09-845-416-10

Query Match 83.3%; Score 1750.8; DB 10; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCATT 60
900 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCATT 959
61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
960 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180

Db 1020 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAGTATTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 240
Db 1080 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGSAAAAATTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGSAAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGSTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGSTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTT-----TTAATGGATCTCCAGATCA 474
Db 1320 TAGCATGGAAAAACAAAGCAATTTTACATAGAACTCATAGATTACTGCAACAGTTCCTCCCT 1379
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGA 534
Db 1380 GGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACA 1439
QY 535 GGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGT 594
Db 1440 GGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAA 1499
QY 595 GCTTCAAGAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGT 654
Db 1500 ACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGA 1559
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCAACCAACTGCTGCTTTTGGAAAGA 702
Db 1560 TGAATAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCTCTGTTACA 1619
QY 703 ACAACTTAAG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAAGCCG 759
Db 1620 AAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAAGTCTCTCAACAT 1679
QY 760 CTGGGTTCTTTTACAAGACAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGA 819
Db 1680 TAGGTCCCATTTTGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGA 1739
QY 820 ACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGG 879
Db 1740 ACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGG 1799
QY 880 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 939
Db 1800 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 1859
QY 940 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 999
Db 1860 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
QY 1000 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCCTCCTGAGGAGAG 1059
Db 1920 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCCTCCTGAGGAGAG 1979
QY 1060 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 1119
Db 1980 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2039
QY 1120 AAAATTGAACCTCGCTGACTCGGAGAGAAAAATAGATGAGACCCCTTGAAAAGACT 1179
Db 2040 AAAATTGAACCTCGCTGACTCGGAGAGAAAAATAGATGAGACCCCTTGAAAAGACT 2099
QY 1180 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 1239
Db 2100 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2159

1240 CAAGGATCTTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 1299
2160 CAAGGATCTTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2219
1300 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 1359
2220 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 2279
1360 TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 1419
2280 TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 2339
1420 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 1479
2340 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 2399
1480 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 1539
2400 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 2459
1540 TGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1599
2460 TGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2519
1600 CGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACCGACTTTTAGC 1659
2520 CGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACCGACTTTTAGC 2579
1660 TGACCTGAATATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
2580 TGACCTGAATATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639
1720 GAAAGCCCTTTGCTTGGATCTCTTGGGACCTGTGAGCTGCTGATGCTTGGACCGCA 1779
2640 GAAAGCCCTTTGCTTGGATCTCTTGGGACCTGTGAGCTGCTGATGCTTGGACCGCA 2699
1780 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAATTAATTTGTTGACCCAC 1839
2700 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAATTAATTTGTTGACCCAC 2759
1840 TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTGTCACAGTCCCTCTCTGCGTGA 1899
2760 TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTGTCACAGTCCCTCTCTGCGTGA 2819
1900 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1959
2820 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 2879
1960 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGACAAGTACAG 2019
2880 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGACAAGTACAG 2939
2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCCT 2079
2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCCT 2999
2080 CCTTCTGCATGATTTCTATCCAA 2101
3000 CCTTCTGCATGATTTCTATCCAA 3021

ESULT 5
S-09-845-416-30
Sequence 30, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 4498
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-30

Query Match 83.3%; Score 1750.8; DB 10; Length 4498;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1717 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTTCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 300
Db 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAACCAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAATCA 474
Db 2077 TAGCATGGAACCAAGCAATTTACATAGAGTTCTATAGAACTCATAGATTCTGCAACAGTTCCCCCT 2136
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGA 534
Db 2137 GGAACCTGGAAGAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCAATGTCTCTACA 2196
QY 535 GGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCACTACAACACATAAGGT 594
Db 2197 GGAATGCTACCCGTAAAGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAGAGAGCTGATGA 2256
QY 595 GCTTCAAGAAGATCTAGAACAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGT 654
Db 2257 ACAATGGCAAGACCTTCAAGGTGAAATTTGAAGCTCACACAGATGTTATCACAACTGGA 2316
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGA 702
Db 2317 TGAACACAGCCAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGATGATGCTGTTACA 2376
QY 703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCG 759
Db 2377 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAAGTCTCTCAACAT 2436
QY 760 CTGGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAGGA 819
Db 2437 TAGTCTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAGGA 2496
QY 820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 879
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 2556

Y 880 CGACTTTCCAGCAGTTTCAAGACGACATGTACATAGGCGCTTCAAGAGGGAATTGAA 939
b |||||||
2557 CGACTTTCCAGCAGTTTCAAGACGACATGTACATAGGCGCTTCAAGAGGGAATTGAA 2616
Y 940 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCA 999
b |||||||
2617 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCA 2676
Y 1000 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 1059
b |||||||
2677 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 2736
Y 1060 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 1119
b |||||||
2737 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2796
Y 1120 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 1179
b |||||||
2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2856
Y 1180 CCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGTAT 1239
b |||||||
2857 CCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGTAT 2916
Y 1240 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 1299
b |||||||
2917 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2976
Y 1300 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 1359
b |||||||
2977 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 3036
Y 1360 TGACCTTGCTCGCCACGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 1419
b |||||||
3037 TGACCTTGCTCGCCACGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 3096
Y 1420 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGCTCGAGGACCGAGTCAG 1479
b |||||||
3097 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGCTCGAGGACCGAGTCAG 3156
Y 1480 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACACTTTCTTTCCACGTC 1539
b |||||||
3157 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACACTTTCTTTCCACGTC 3216
Y 1540 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 1599
b |||||||
3217 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 3276
Y 1600 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 1659
b |||||||
3277 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 3336
Y 1660 TGACCTGAATAATGTACAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
b |||||||
3337 TGACCTGAATAATGTACAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396
Y 1720 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGACCCAGCA 1779
b |||||||
3397 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGACCCAGCA 3456
Y 1780 CAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTTGTTGACCCAC 1839
b |||||||
3457 CAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTTGTTGACCCAC 3516
Y 1840 TATTTATGACCGCCTGGAGCAAGAGCACAACATTTGGTCAACGCTCCTCTCTGCGTGA 1899
b |||||||
3517 TATTTATGACCGCCTGGAGCAAGAGCACAACATTTGGTCAACGCTCCTCTCTGCGTGA 3576
Y 1900 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1959
b |||||||
3577 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3636

QY 1960 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAG 2019
Db |||||||
3637 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAG 3696
QY 2020 AFACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 2079
Db |||||||
3697 AFACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 3756
QY 2080 CCTTCTGCATGATTCTATCCAA 2101
Db |||||||
3757 CCTTCTGCATGATTCTATCCAA 3778
RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
Query Match 83.0%; Score 1743; DB 10; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCCCTACACACAGCGCTGCTTATGTCAACACCTCTGACCCCTACCGAGCCCAT 60
Db |||||||
900 GAGCTATGCCCTACACACAGCGCTGCTTATGTCAACACCTCTGACCCCTACCGAGCCCAT 959
QY 61 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db |||||||
960 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
Db |||||||
1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||||
1080 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTTGG 300
Db |||||||
1140 CCAGTTTCATACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTTGG 1199
QY 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 360
Db |||||||
1200 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db |||||||
1260 AACTGAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 480
Db |||||||
1320 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db |||||||
1380 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1439

LENGTH: 4825									
TYPE: DNA									
ORGANISM: Homo sapiens									
3-09-845-416-29									
Query Match 83.0%; Score 1743; DB 10; Length 4825;									
Best Local Similarity 85.8%; Pred.No. 0;									
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;									
/	1	GAGCTATGCTACACACAGGCTGCTTATGTCA	CCACCTCTGACCCCTACACGGAGCCCA	TT 60					
/	1657	GAGCTATGCTACACACAGGCTGCTTATGTCA	CCACCTCTGACCCCTACACGGAGCCCA	TT 1716					
/	61	TCCTTCACAGCATTTGGAAGCTCCTGAAGACA	CAAGTCAATTGGCAGTTCATTGATGGAGAG	120					
/	1717	TCCTTCACAGCATTTGGAAGCTCCTGAAGACA	CAAGTCAATTGGCAGTTCATTGATGGAGAG	1776					
/	121	TGAAGTAAACCTGGACCGTTATCAAAACAGCT	TTTAGAAGAAGTATTATCGTGGCTTCTTTC	180					
/	1777	TGAAGTAAACCTGGACCGTTATCAAAACAGCT	TTTAGAAGAAGTATTATCGTGGCTTCTTTC	1836					
/	181	TGCTGAGGACACATTTGCAAGCACAAAGGAGA	GATTTTCTAATGATGTGGAAGTGGTGAAGA	240					
/	1837	TGCTGAGGACACATTTGCAAGCACAAAGGAGA	GATTTTCTAATGATGTGGAAGTGGTGAAGA	1896					
/	241	CCAGTTTCATACTCATGAGGGGTACATGATGG	ATTTTGACAGCCCATCAGGGCCGGGTGG	300					
b	1897	CCAGTTTCATACTCATGAGGGGTACATGATGG	ATTTTGACAGCCCATCAGGGCCGGGTGG	1956					
/	301	TAATATTCTACAATTGGGAAGTAAGCTGATT	GGAAACAGGAAAATTATCAGAAGATGAAGA	360					
b	1957	TAATATTCTACAATTGGGAAGTAAGCTGATT	GGAAACAGGAAAATTATCAGAAGATGAAGA	2016					
/	361	AACGTGAAGTACAAGACAGATGAATCTCCTAA	ATTCAAGATGGGAATGCCTCAGGGTAGC	420					
b	2017	AACGTGAAGTACAAGACAGATGAATCTCCTAA	ATTCAAGATGGGAATGCCTCAGGGTAGC	2076					
/	421	TAGCATGGAAAAACAAGCAATTTACATAGAGT	TTTAAATGGATCTCCAGAATCAGAAACT	480					
b	2077	TAGCATGGAAAAACAAGCAATTTACATAGAGT	TTTAAATGGATCTCCAGAATCAGAAACT	2136					
/	481	GAAAGAGTTGAATGACTGGCTAACAAAAACA	GAAAGAAATGGAGGAAGA 540						
b	2137	GAAAGAGTTGAATGACTGGCTAACAAAAACA	GAAAGAAATGGAGGAAGA 2196						
/	541	GCCTCTTGGACCTGATCTTGAAGACCTTAA	AAACGCCAAGTACAACAATAAGGTGCTTCA	600					
b	2197	GCCTCTTGGACCTGATCTTGAAGACCTTAA	AAACGCCAAGTACAACAATAAGGTGCTTCA	2256					
/	601	AGAAGATCTAGAACAAAGAACAAAGTCAAGG	TCAATTCTCTCACTCACATGGTGGTGTAGT	660					
b	2257	AGAAGATCTAGAACAAAGAACAAAGTCAAGG	TCAATTCTCTCACTCACATGGTGGTGTAGT	2316					
/	661	TGATGAATCTAGTGGAGATCACGCAACTGCT	GTCTTTGGAAGAACAACTTAAGGTATTGGG	720					
b	2317	TGATGAATCTAGTGGAGATCACGCAACTGCT	GTCTTTGGAAGAACAACTTAAGGTATTGGG	2376					
/	721	AGATCGATGGCAACATCTGTAGATGGACAGA	AGACCGCTGGGTTCTTTTACAAGA --- 777						
b	2377	AGATCGATGGCAACATCTGTAGATGGACAGA	AGACCGCTGGGTTCTTTTACAAGACAC 2436						
/	778	-----	-----	777					
b	2437	TCATAGATTACTGCAACAGTTCCCTGGACCT	GGAAAAGTTTCTTTCCTGGCTTACAGA 2496						
/	778	-----	-----	777					
b	2497	AGCTGAAACAACTGCCAATGTCTTACAGGAT	GCTACCCGTAAGGAAAGGCTCCTAGAGA 2556						
/	778	-----	-----	777					
b	2557	CTCCAAGGGAGTAAAGAGCTGATGAAACA	ATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616						

Qy	778	-----	-----	777
Db	2617	TCACACAGATGTTTATCAACAACCTGGATGA	AAAAACAGCCAAAAAATCCTGAGATCCCTGGA	2676
Qy	778	-----	-----	777
Db	2677	AGGTTCCGATGATGCAGTCTCTGTTACAAA	GACGTTTGGATAACATGAACCTTCAAGTGGAG	2736
Qy	778	-----	-----	792
Db	2737	TGAACCTTCGGAAAAAGTCTCTCAACATTAG	GTCCCATTGGAAGCCAGTTTCTGACCCAGTG	2796
Qy	793	GAAGCGTCTGCACCTTTCTCTGCAGGAACT	TTCTGGTGTGGCTACAGCTGAAAAGATGATGA	852
Db	2797	GAAGCGTCTGCACCTTTCTCTGCAGGAACT	TTCTGGTGTGGCTACAGCTGAAAAGATGATGA	2856
Qy	853	ATTAAGCCGGCAGCACCTATTGGAGGCGAC	TTTCCAGCAGTTTCAGAAAGCAGAACCGATGT	912
Db	2857	ATTAAGCCGGCAGCACCTATTGGAGGCGAC	TTTCCAGCAGTTTCAGAAAGCAGAACCGATGT	2916
Qy	913	ACATAGGGCCCTTCAAGAGGGGAATTGAAA	AACTAAAGAACCTGTAAATCATGACTCTTTGA	972
Db	2917	ACATAGGGCCCTTCAAGAGGGGAATTGAAA	AACTAAAGAACCTGTAAATCATGACTCTTTGA	2976
Qy	973	GACTGTACGAATATTCTGACAGAGCAGCCT	TTTGGAAAGGACTAGAGAAACTCTACCAGGA	1032
Db	2977	GACTGTACGAATATTCTGACAGAGCAGCCT	TTTGGAAAGGACTAGAGAAACTCTACCAGGA	3036
Qy	1033	GGCCAGAGAGCTGCCTCTGAGGAGAGAGCC	CAAGATGTCACTCGGCTTCTACGAAAGCA	1092
Db	3037	GGCCAGAGAGCTGCCTCTGAGGAGAGAGCC	CAAGATGTCACTCGGCTTCTACGAAAGCA	3096
Qy	1093	GGCTGAGGAGGTCAATACTGAGTGGGAAA	AAATTGAACTGCACCTCCGCTGACTGGCAGAG	1152
Db	3097	GGCTGAGGAGGTCAATACTGAGTGGGAAA	AAATTGAACTGCACCTCCGCTGACTGGCAGAG	3156
Qy	1153	AAAAATAGATGAGACCCCTTGAAGACTCC	AGAACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA	1212
Db	3157	AAAAATAGATGAGACCCCTTGAAGACTCC	AGAACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA	3216
Qy	1213	CCTCAAGCTGCGCCAAAGCTGAGGTGATCA	AGGGATCCTGGCAGCCCGTGGGCGATCTCCT	1272
Db	3217	CCTCAAGCTGCGCCAAAGCTGAGGTGATCA	AGGGATCCTGGCAGCCCGTGGGCGATCTCCT	3276
Qy	1273	CATTGACTCTCTCCAAGATCACCTCGAGA	AAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC	1332
Db	3277	CATTGACTCTCTCCAAGATCACCTCGAGA	AAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC	3336
Qy	1333	TCGAAAAGAGAACGCTGAGCCACAGTCAAT	GACCTTGTCTGCCAGCTTTACCACCTTTGGGCA	1392
Db	3337	TCGAAAAGAGAACGCTGAGCCACAGTCAAT	GACCTTGTCTGCCAGCTTTACCACCTTTGGGCA	3396
Qy	1393	TCAGCTCTCACCGTATAACCTCAGCACTCT	GGAAGACCTGAAACACCCAGATGGAAGCTTCT	1452
Db	3397	TCAGCTCTCACCGTATAACCTCAGCACTCT	GGAAGACCTGAAACACCCAGATGGAAGCTTCT	3456
Qy	1453	GCAGGTGGCCGTCGAGGACCGAGTCAGGC	AGTCAAGCCACAGCCACAGGGACTTTGGTCC	1512
Db	3457	GCAGGTGGCCGTCGAGGACCGAGTCAGGC	AGTCAAGCCACAGCCACAGGGACTTTGGTCC	3516
Qy	1513	AGCATCTCAGCACTTTTCTCCACGTCGTCT	GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	1572
Db	3517	AGCATCTCAGCACTTTTCTCCACGTCGTCT	GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	3576
Qy	1573	AAACAAAGTGCCCTACTATATCAACCAGAG	ACTCAAACAACTTGTCTGGGACCATCCCAA	1632
Db	3577	AAACAAAGTGCCCTACTATATCAACCAGAG	ACTCAAACAACTTGTCTGGGACCATCCCAA	3636
Qy	1633	AATGACAGAGCTCTACCAGTCTTTTAGCTG	ACCTGAATAATGTGAGATTCTCAGCTTATAG	1692
Db	3637	AATGACAGAGCTCTACCAGTCTTTTAGCTG	ACCTGAATAATGTGAGATTCTCAGCTTATAG	3696
Qy	1693	GACTGCCATGAAAACCTCCGAAGACTGCAG	AAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	1752

||||| 3697 GACTGCCATGAAACTCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTTTGAGCCTGTC 3756
||||| 1753 AGCTGCATGTGATGCCCTGGACCAACAACTCAAGCAAAATGACCAGCCCATGGATAT 1812
||||| 3757 AGCTGCATGTGATGCCCTGGACCAACAACTCAAGCAAAATGACCAGCCCATGGATAT 3816
||||| 1813 CCTGCAGATTATTAATGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCACAA 1872
||||| 3817 CCTGCAGATTATTAATGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCACAA 3876
||||| 1873 TTTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGA 1932
||||| 3877 TTTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGA 3936
||||| 1933 TACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTG 1992
||||| 3937 TACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTG 3996
||||| 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
||||| 3997 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4056
||||| 2053 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGTCATGATTTCTATCCAA 2101
||||| 4057 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGTCATGATTTCTATCCAA 4105

RESULT 8

JS-09-845-416-35
Sequence 35, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 4848
TYPE: DNA
ORGANISM: Homo sapiens
JS-09-845-416-35

Query Match 83.0%; Score 1743; DB 10; Length 4848;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
2Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
Db ||||| 1680 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 1739
2Y 61 TCCTTCACAGCATTTGGAAGCTCTGGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 120
Db ||||| 1740 TCCTTCACAGCATTTGGAAGCTCTGGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 1799
2Y 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
Db ||||| 1800 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 1859
2Y 181 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db ||||| 1860 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1919
2Y 241 CCAGTTTCACTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 300
Db ||||| 1920 CCAGTTTCACTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 1979

QY 301 TAATATTCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db ||||| 1980 TAATATTCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2039
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
Db ||||| 2040 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 2099
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db ||||| 2100 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2159
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
Db ||||| 2160 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 2219
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCGCAAGTACAAACAACATAAGGTGCTTCA 600
Db ||||| 2220 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCGCAAGTACAAACAACATAAGGTGCTTCA 2279
QY 601 AGAAGATCTAGAAACAAGAAACAAGTCAGGTCAGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db ||||| 2280 AGAAGATCTAGAAACAAGAAACAAGTCAGGTCAGGTCAATTTCTCACTCACATGGTGGTAGT 2339
QY 661 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTGGAAGAAACAACATTAAGGTATTGGG 720
Db ||||| 2340 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTGGAAGAAACAACATTAAGGTATTGGG 2399
QY 721 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGA--- 777
Db ||||| 2400 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC 2459
QY 778 ----- 777
Db ||||| 2460 TCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAGAAAGTTTCTTGCCTGGCTTACAGA 2519
QY 778 ----- 777
Db ||||| 2520 AGCTGAAACAACCTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 2579
QY 778 ----- 777
Db ||||| 2580 CTCCAAGGGAGTAAAGAGCTGTATGATAACAATGGCAAGACCTCCAAGGTGAATTAAGC 2639
QY 778 ----- 777
Db ||||| 2640 TCACACAGATGTTTATCAACACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2699
QY 778 ----- 777
Db ||||| 2700 AGGTTCCGATGATGCAGTCCTGTATCAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2759
QY 778 -----CAGTTCTGACCCAGTG 792
Db ||||| 2760 TGAACCTTCGGAAAAAGTCTCTCAACATTTAGTCCCATTTGGAAGCCAGTTCTGACCCAGTG 2819
QY 793 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db ||||| 2820 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2879
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 912
Db ||||| 2880 ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 2939
QY 913 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGAGTACTCTTGA 972
Db ||||| 2940 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGAGTACTCTTGA 2999
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAAGACTAGAGAAACTCTACCAGGA 1032
Db ||||| 3000 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAAGACTAGAGAAACTCTACCAGGA 3059
QY 1033 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGACA 1092

b 3060 G C C C A G A G A G C T G C C T C C T G A G G A G A G C C C A G A A T G T C A C T C G G C T T C T A C G A A A G C A 3119
y 1093 G G C T G A G G A G G T C A A T A C T A G T G G G A A A A A T T G A A C C T G C A C T C C G C T G A C T G G C A G A G 1152
b 3120 G G C T G A G G A G G T C A A T A C T A G T G G G A A A A A T T G A A C C T G C A C T C C G C T G A C T G G C A G A G 3179
y 1153 A A A A A T A G A T G A G A C C C T T G A A A G A C T C C A G A A A C T T C A A G A G C C C A C G G A T G A G T G G A 1212
b 3180 A A A A A T A G A T G A G A C C C T T G A A A G A C T C C A G A A A C T T C A A G A G C C C A C G G A T G A G T G G A 3239
y 1213 C C T C A A G C T G C C C A A G C T G A G G T G A T C A A G G G A T C C T G G C A G C C C G T G G C G A T C T C C T 1272
b 3240 C C T C A A G C T G C C C A A G C T G A G G T G A T C A A G G G A T C C T G G C A G C C C G T G G C G A T C T C C T 3299
y 1273 C A T T G A C T C T C T C C A A G A T C A C C T C G A G A A G T C A A G G C A C T T C G A G G A G A A A T T G C G C C 1332
b 3300 C A T T G A C T C T C T C C A A G A T C A C C T C G A G A A G T C A A G G C A C T T C G A G G A G A A A T T G C G C C 3359
y 1333 T C T G A A A G A G A A C G T G A G C C C A C G T C A A T G A C C T T G C T C G C C A G C T T A C C A C T T T G G G C A T 1392
b 3360 T C T G A A A G A G A A C G T G A G C C C A C G T C A A T G A C C T T G C T C G C C A G C T T A C C A C T T T G G G C A T 3419
y 1393 T C A G C T C T C A C C G T A A A C C T C A G C A C T C T G G A A G A C C T G A A C A C C A G A T G G A A G C T T C T 1452
b 3420 T C A G C T C T C A C C G T A T A A C C T C A G C A C T C T G G A A G A C C T G A A C A C C A G A T G G A A G C T T C T 3479
y 1453 G C A G G T G G C C G T C G A G A C C G A G T C A G G C A G C T G C A T G A A G C C C A C A G G A C T T T G G T C C 1512
b 3480 G C A G G T G G C C G T C G A G A C C G A G T C A G G C A G C T G C A T G A A G C C C A C A G G A C T T T G G T C C 3539
y 1513 A G C A T C T C A G C A C T T T C T T C C A C G T C T G T C C A G G G T C C C T G G A G A G A G C C A T C T C G C C 1572
b 3540 A G C A T C T C A G C A C T T T C T T C C A C G T C T G T C C A G G G T C C C T G G A G A G A G C C A T C T C G C C 3599
y 1573 A A A C A A A G T G C C C T A C T A T A T C A A C C C A G A G A C T C A A A C A A C T T G C T G G A C C A T C C C A A 1632
b 3600 A A A C A A A G T G C C C T A C T A T A T C A A C C C A G A G A C T C A A A C A A C T T G C T G G A C C A T C C C A A 3659
y 1633 A A T G A C A G A G C T C T A C C A G T C T T T A G C T G A C C T G A A T A A T G T C A G A T T C T C A G C T T A T A G 1692
b 3660 A A T G A C A G A G C T C T A C C A G T C T T T A G C T G A C C T G A A T A A T G T C A G A T T C T C A G C T T A T A G 3719
y 1693 G A C T G C C A T G A A A C T C C G A A G A C T G C A A A G G C C C T T T G C T T G G A T C T C T T G A G C C T G T C 1752
b 3720 G A C T G C C A T G A A A C T C C G A A G A C T G C A A A G G C C C T T T G C T T G G A T C T C T T G A G C C T G T C 3779
y 1753 A G C T G C A T G T A T G C C T T G G A C C A G A A C C T C A A G C A A A A T G A C C A G C C C A T G G A T A T 1812
b 3780 A G C T G C A T G T A T G C C T T G G A C C A G A C A A C C T C A A G C A A A A T G A C C A G C C C A T G G A T A T 3839
y 1813 C C T G C A G A T T A T T A A T T G T T T G A C C A C T A T T A T A G C C G C C T G G A G C A A G A G C A A C A A 1872
b 3840 C C T G C A G A T T A T T A A T T G T T T G A C C A C T A T T A T A G C C G C C T G G A G C A A G A G C A A C A A 3899
y 1873 T T T G G T C A A C G T C C C T C T C T G C G T G G A T A T G T C T G A C T G G C T G C T G A A T G T T A T G A 1932
b 3900 T T T G G T C A A C G T C C C T C T C T G C G T G G A T A T G T C T G A A C T G G C T G C T G A A T G T T A T G A 3959
y 1933 T A C G G A C G A A C A G G A G G A T C C G T G T C T G T C T T T T A A A A C T G G C A T C A T T T C C C T G T G 1992
b 3960 T A C G G A C G A A C A G G A G G A T C C G T G T C T G T C T T T T A A A A C T G G C A T C A T T T C C C T G T G 4019
y 1993 T A A A G C A C A T T T G G A A G A C A A G T A C A G A T A C C T T T T C A A G C A A G T G G C A A G T T C A A C A G G 2052
b 4020 T A A A G C A C A T T T G G A A G A C A A G T A C A G A T A C C T T T T C A A G C A A G T G G C A A G T T C A A C A G G 4079
y 2053 A T T T T G T G A C C A G C A G G C T G G G C C T C C T T C T G C A T G A T T C T A T C C A A 2101
b 4080 A T T T T G T G A C C A G C A G G C T G G G C C T C C T T C T G C A T G A T T C T A T C C A A 4128

US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: Del142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 83.0%; Score 1743; DB 10; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

QY 1 GAGCTATGCCTACACACAGCGCTGCTTATGTACCACCTCTGACCCCTACACGAGCCCAT 60
Db 1892 GAGCTATGCCTACACACAGCGCTGCTTATGTACCACCTCTGACCCCTACACGAGCCCAT 1951

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
Db 1952 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 2011

QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 2012 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 2071

QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 2072 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 2132 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 2191

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAATAATATCAGAAGATGAAGA 360
Db 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAATAATATCAGAAGATGAAGA 2251

QY 361 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATCCCTCAGGGTAGC 420
Db 2252 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATCCCTCAGGGTAGC 2311

QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 2312 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2371

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACACGAAAAATGGAGGAAGA 540
Db 2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACACGAAAAATGGAGGAAGA 2431

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAAGCCCAAGTACAAACATAAGGTGCTTCA 600
Db 2432 GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAAGCCCAAGTACAAACATAAGGTGCTTCA 2491

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTAGT 660
Db 2492 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTAGT 2551

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2611

QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTCTTTTACAAGA --- 777

Db 2612 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTACAAGACAC 2671
2Y 778 -----
Db 2672 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTTGGAAAGTTTCTTGCCTGGCTTACAGA 2731
2Y 778 -----
Db 2732 AGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGA 2791
2Y 778 -----
Db 2792 CTCGAAGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 2851
2Y 778 -----
Db 2852 TCACACAGATGTTTATACAAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2911
2Y 778 -----
Db 2912 AGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGATACATGAACCTTCAAGTGGAG 2971
2Y 778 -----
Db 2972 TGAACCTCGAAAAAGTCTCTCAACATTAGGTCCCAATTGGAAGCCAGTTCTGACCAGTG 3031
2Y 793 GAAGGCTCTGCACCTTCTCTGTCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGA 852
3032 GAAGGCTCTGCACCTTCTCTGTCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGA 3091
2Y 853 ATTAAGCCGGCAGGCACCTATTGAGGCGGACCTTCCAGCAGTTTCAGAAGCAGAACGATGT 912
3092 ATTAAGCCGGCAGGCACCTATTGAGGCGGACCTTCCAGCAGTTTCAGAAGCAGAACGATGT 3151
2Y 913 ACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGA 972
3152 ACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGA 3211
2Y 973 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1032
3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 3271
2Y 1033 GCCCAGAGAGCTGCCTCCTTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1092
3272 GCCCAGAGAGCTGCCTCCTTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 3331
2Y 1093 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAG 1152
3332 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAG 3391
2Y 1153 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCGCACGGATGAGCTGGA 1212
3392 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCGCACGGATGAGCTGGA 3451
2Y 1213 CCTCAAGCTGCGCCAAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1272
3452 CCTCAAGCTGCGCCAAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3511
2Y 1273 CATTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCC 1332
3512 CATTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCC 3571
2Y 1333 TCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 1392
3572 TCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 3631
2Y 1393 TCAGCTCTCACCGTATACTCAGCACTCTGGAAGACCTGGAACCAACAGATGGAAGCTTCT 1452
3632 TCAGCTCTCACCGTATACTCAGCACTCTGGAAGACCTGGAACCAACAGATGGAAGCTTCT 3691
2Y 1453 GCAGGTGGCCGTCGAGGACCGGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 1512

Db 3692 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGACTTTGGTCC 3751
2Y 1513 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 1572
Db 3752 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 3811
2Y 1573 AAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCAA 1632
3812 AAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCAA 3871
2Y 1633 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTAGATTCTCAGCTTATAG 1692
3872 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTAGATTCTCAGCTTATAG 3931
2Y 1693 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
3932 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3991
2Y 1753 AGCTGCATGTGATGCCCTTGACCCAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATAT 1812
3992 AGCTGCATGTGATGCCCTTGACCCAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATAT 4051
2Y 1813 CCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAA 1872
4052 CCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAA 4111
2Y 1873 TTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGA 1932
4112 TTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGA 4171
2Y 1933 TACGGACGAAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTG 1992
4172 TACGGACGAAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTG 4231
2Y 1993 TAAAGCACATTTTGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGG 2052
4232 TAAAGCACATTTTGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGG 4291
2Y 2053 ATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
4292 ATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 4340

RESULT 10
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 78.1%; Score 1641; DB 10; Length 4414;
Best local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;
2Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1716
2Y 61 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120

b 1717 TCCTTACAGCATTTGGAAGCTCCTGAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1776
y 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTTCTTC 180
b 1777 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTTCTTC 1836
y 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAAGTGGTGAAGA 240
b 1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAAGTGGTGAAGA 1896
y 241 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG 300
b 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG 1956
y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAATAATTATCAGAAGATGAAGA 360
b 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAATAATTATCAGAAGATGAAGA 2016
y 361 AACTGAAGTACAAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGCTAGC 420
b 2017 AACTGAAGTACAAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGCTAGC 2076
y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
b 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
y 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAAAGAAATGGAGGAAGA 540
b 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAAAGAAATGGAGGAAGA 2196
y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 600
b 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 2256
y 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
b 2257 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 2316
y 661 TGATGAATCTAGTGGAGATCACGCCAAGTCTGCTTTTGGAAAGAACAACTTAAGTATTGGG 720
b 2317 TGATGAATCTAGTGGAGATCACGCCAAGTCTGCTTTTGGAAAGAACAACTTAAGTATTGGG 2376
y 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAG 780
b 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 2436
y 781 TTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGAACTTCTGGTGTGGTACAGCT 840
b 2437 CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTCAGA 2496
y 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA 900
b 2497 AAAAGAAGATGCAGTGAACAAGATTTCACAACAACTGGCTTTTAAAGATCAAAATGAAATGTT 2556
y 901 GCAGAACGATGPACATAGGGCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAAATCAT 960
b 2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2616
y 961 GAGTACTCTTGAGACTGTACGAATAATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAA 1020
b 2617 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG 2662
y 1021 ACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1080
b 2663 AATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGG-- 2720
y 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGC 1140
b 2721 -----ATAATTTAGTCCAAAAACTTGAA----- 2743
y 1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1200

RESULT 11
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

Db 2744 -----AAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCAC 2793
QY 1201 GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
Db 2794 GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2853
QY 1261 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG 1320
Db 2854 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG 2913
QY 1321 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTAC 1380
Db 2914 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTAC 2973
QY 1381 CACTTTGGGCATTTCAAGCTCTACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
Db 2974 CACTTTGGGCATTTCAAGCTCTACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 3033
QY 1441 ATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCACAG 1500
Db 3034 ATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCACAG 3093
QY 1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAG 1560
Db 3094 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAG 3153
QY 1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTG 1620
Db 3154 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTG 3213
QY 1621 GGACCATCCAAAAATGACAGAGCTCTACCAAGTCTTAGCTGACCTGAATAATGTGAGATT 1680
Db 3214 GGACCATCCAAAAATGACAGAGCTCTACCAAGTCTTAGCTGACCTGAATAATGTGAGATT 3273
QY 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 1740
Db 3274 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 3333
QY 1741 CTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCACCAACCTCAAGCAAAATGACCA 1800
Db 3334 CTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCACCAACCTCAAGCAAAATGACCA 3393
QY 1801 GCCATGGATATCCTGCAGATTATTAATTGTTGACCACATATTTATGACCCGCTGGAGCA 1860
Db 3394 GCCATGGATATCCTGCAGATTATTAATTGTTGACCACATATTTATGACCCGCTGGAGCA 3453
QY 1861 AGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCT 1920
Db 3454 AGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCT 3513
QY 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGTCTCTGCTTTTAAACCTGGCAT 1980
Db 3514 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGTCTCTGCTTTTAAACCTGGCAT 3573
QY 1981 CATTTCCCTGTGTAAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db 3574 CATTTCCCTGTGTAAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 3633
QY 2041 AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTTCTATCCA 2100
Db 3634 AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTTCTATCCA 3693
QY 2101 A 2101
Db 3694 A 3694

TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142

CURRENT APPLICATION NUMBER: US/09/845,416

CURRENT FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 60/200,777

PRIOR FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 14

LENGTH: 3446

TYPE: DNA

ORGANISM: Homo sapiens

IS-09-845-416-14

Query Match	77.5%;	Score 1629;	DB 10;	Length 3446;
Best Local Similarity	87.9%;	Pred. No. 0;		
Matches 1847;	Conservative	0;	Mismatches 190;	Indels 64; Gaps 4;
y	1	GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAATT	60	
b	900	GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAATT	959	
y	61	TCCTTCACAGCATTTGGAGCTCCTGAAGACAACTGATTTGGCAGTTCATTGATGGAGAG	120	
b	960	TCCTTCACAGCATTTGGAGCTCCTGAAGACAACTGATTTGGCAGTTCATTGATGGAGAG	1019	
y	121	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC	180	
b	1020	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC	1079	
y	181	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	240	
b	1080	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA	1139	
y	241	CCAGTTTCATACACTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGTTGG	300	
b	1140	CCAGTTTCATACACTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGTTGG	1199	
y	301	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGGAACAGGAAAAATTATCAGAAGATGAAGA	360	
b	1200	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGGAACAGGAAAAATTATCAGAAGATGAAGA	1259	
y	361	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC	420	
b	1260	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC	1319	
y	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT	480	
b	1320	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAACT	1378	
y	481	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAGGAAAAATGGAGGAAGA	540	
b	1379	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAGGAAAAATGGAGGAAGA	1438	
y	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACAATAAGGTGCTTCA	600	
b	1439	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACAATAAGGTGCTTCA	1498	
y	601	AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT	660	
b	1499	AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT	1558	
y	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG	720	
b	1559	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG	1618	
y	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGCCGCTGGGTTCTTTTACAAGACAG	780	
b	1619	AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGCCGCTGGGTTCTTTTACAAGACAT	1678	
y	781	TTCTGACCACTGGGAAGCGTCTGCACCTTTCTCTGCAAGAACTTCTGGTGGCTACAGCT	840	

Db	1679	CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGATGGCTTTTCAGA	1738
QY	841	GAAAGATGATGAATTAAAGCGGCGAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA	900
Db	1739	AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT	1798
QY	901	GCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCAT	960
Db	1799	ATCAAGTCTTCAAAAACCTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT	1858
QY	961	GAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAGGACTAGAGAA	1020
Db	1859	GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG	1904
QY	1021	ACTCTACCAGGAGCCCGACAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCT	1080
Db	1905	AATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGA-----	1942
QY	1081	TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGC	1140
Db	1943	-----TAACTTTGCCCGGTGTGGGATAATTTTAGTCCAAAAACTTGAA-----	1985
QY	1141	TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAGACTCCAGGAACCTTCAAGAGGCCAC	1200
Db	1986	-----AAGACTACAGCACAGACCCCTTTGAAGACTCCAGGAACCTTCAAGAGGCCAC	2035
QY	1201	GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTTGGCAGCCCGT	1260
Db	2036	GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTTGGCAGCCCGT	2095
QY	1261	GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG	1320
Db	2096	GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG	2155
QY	1321	AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC	1380
Db	2156	AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC	2215
QY	1381	CACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG	1440
Db	2216	CACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG	2275
QY	1441	ATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAAGGCACTGCATGAAGCCCAACAG	1500
Db	2276	ATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAAGGCACTGCATGAAGCCCAACAG	2335
QY	1501	GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAG	1560
Db	2336	GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAG	2395
QY	1561	AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGCTG	1620
Db	2396	AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGCTG	2455
QY	1621	GGACCATCCCAAAATGACAGAGTCTACCAAGTCTTTAGCTGACCTGAATATGTCAAGTT	1680
Db	2456	GGACCATCCCAAAATGACAGAGTCTACCAAGTCTTTAGCTGACCTGAATATGTCAAGTT	2515
QY	1681	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCT	1740
Db	2516	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCT	2575
QY	1741	CTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCA	1800
Db	2576	CTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCA	2635
QY	1801	GCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	1860
Db	2636	GCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	2695
QY	1861	AGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACCTGGCTGCT	1920
Db	2696	AGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACCTGGCTGCT	2755

1921 GAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTGTCCTTTTAAACTGGCAT 1980
2756 GAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTGTCCTTTTAAACTGGCAT 2815
1981 CATTTCCCTGTGTAAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
2816 CATTTCCCTGTGTAAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2875
2041 AAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCA 2100
2876 AAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCA 2935
2101 A 2101
2936 A 2936

RESULT 12
Sequence 41, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 5462
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
S-10-149-736-41

Query Match 77.5%; Score 1629; DB 15; Length 5462;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;

1 GAGCTATGCCTACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
1099 GAGCTATGCCTACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158
61 TCCTTCACAGCATTTGGAAGCTCCTGAAACAAAGTCAATTTGGCAGTTTCATTGATGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAACAAAGTCAATTTGGCAGTTTCATTGATGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGTATATCGTGGCTTCTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGTATATCGTGGCTTCTTTC 1278
181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCATACCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
1339 CCAGTTTCATACCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1398
301 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAAACAGGAAATATCAGAAGATGAAGA 360
1399 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAAACAGGAAATATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420

Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT----- 454
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTGACCACTATTGGAGCCCTCTCC 1578
QY 455 ----- 454
Db 1579 TACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAA 1638
QY 455 -----TTAATGGATCTCCAGAAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCTTGTGATGTTGGAGCATAGATTACTGCAACAGTTCCTCCCTCGA 1698
QY 478 ACTGAAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAGGAAAAATGGAGGA 537
Db 1699 CCTGAAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGA 1758
QY 538 AGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAAGGTGCT 597
Db 1759 TGCTACCCGTAAAGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACA 1818
QY 598 TCAAGAAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGCTGGTGGT 657
Db 1819 ATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGA 1878
QY 658 AGTTGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACA 705
Db 1879 AAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTTTACAAAG 1938
QY 706 ACTTAAGG--TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTG 762
Db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAG 1998
QY 763 GGTTCCTTTTACAAGACAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACT 822
Db 1999 GTCCCATTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACT 2058
QY 823 TCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCAACCTATTGGAGGCGA 882
Db 2059 TCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGGCGA 2118
QY 883 CTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAAAC 942
Db 2119 CTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAAAC 2178
QY 943 TAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 1002
Db 2179 TAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 2238
QY 1003 TTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGC 1062
Db 2239 TTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGC 2298
QY 1063 CCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1122
Db 2299 CCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 2358
QY 1123 ATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCA 1182
Db 2359 ATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCA 2418
QY 1183 GGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAA 1242
Db 2419 GGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAA 2478
QY 1243 GGGATCCTGGCAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAA 1302
Db 2479 GGGATCCTGGCAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAA 2538
QY 1303 AGTCAAGGCACCTTCGAGGAGAAATTCGCTCTCAATTGACTCTCTCCAAGATCACCTCGAGAA 1362
Db 2539 AGTCAAGGCACCTTCGAGGAGAAATTCGCTCTCAATTGACTCTCTCCAAGATCACCTCGAGAA 2598

Y 1363 CCTTGTGCGCAGCTTACCACTTTGGGCATTGAGCTCTCAGCTTACCGTATAACCTCAGCACTCT 1422
b |||||||
Y 2599 CCTTGTGCGCAGCTTACCACTTTGGGCATTGAGCTCTCAGCTTACCGTATAACCTCAGCACTCT 2658
b |||||||
Y 1423 GGAAGACCTGAACACACAGATGGAAGCTTCTGCGAGTGGCGTGGAGGACCGAGTCAGGCA 1482
b |||||||
Y 2659 GGAAGACCTGAACACACAGATGGAAGCTTCTGCGAGTGGCGTGGAGGACCGAGTCAGGCA 2718
b |||||||
Y 1483 GCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGT 1542
b |||||||
Y 2719 GCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGT 2778
b |||||||
Y 1543 CCAGGGTCCCTGGGAGAGCCCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACACGCA 1602
b |||||||
Y 2779 CCAGGGTCCCTGGGAGAGCCCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACACGCA 2838
b |||||||
Y 1603 GACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTTAGCTGA 1662
b |||||||
Y 2839 GACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTTAGCTGA 2898
b |||||||
Y 1663 CCTGAATAANTGTAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 1722
b |||||||
Y 2899 CCTGAATAANTGTAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 2958
b |||||||
Y 1723 GGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGAGCTGATGATGCTTGGACCAACAA 1782
b |||||||
Y 2959 GGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGAGCTGATGATGCTTGGACCAACAA 3018
b |||||||
Y 1783 CCTCAAGCAAAATGACAGAGCCCATGGATATCTGAGATATCTGAGATTAATTTGTTGACCACTAT 1842
b |||||||
Y 3019 CCTCAAGCAAAATGACAGAGCCCATGGATATCTGAGATTAATTTGTTGACCACTAT 3078
b |||||||
Y 1843 TTATGACCGCCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATAT 1902
b |||||||
Y 3079 TTATGACCGCCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATAT 3138
b |||||||
Y 1903 GTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGTCT 1962
b |||||||
Y 3139 GTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGTCT 3198
b |||||||
Y 1963 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGCAAGTACAGATA 2022
b |||||||
Y 3199 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGCAAGTACAGATA 3258
b |||||||
Y 2023 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGGAGGCTGGCCCTCT 2082
b |||||||
Y 3259 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGGAGGCTGGCCCTCT 3318
b |||||||
Y 2083 TCTGCATGATTTCTATCCAA 2101
b |||||||
Y 3319 TCTGCATGATTTCTATCCAA 3337
b |||||||

RESULT 13
US-10-149-736-39
Sequence 39, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 5417

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-39
Query Match 76.7%; Score 1611.6; DB 15; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGAGCCCAATT 60
Db |||||||
QY 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGAGCCCAATT 1158
Db |||||||
QY 61 TCCTTTCACAGCATTTGGAGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db |||||||
QY 1159 TCCTTTCACAGCATTTGGAGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
Db |||||||
QY 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCCTTC 180
Db |||||||
QY 1219 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCCTTC 1278
Db |||||||
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||||
QY 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
Db |||||||
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGAGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db |||||||
QY 1339 CCAGTTTCATCTCATGAGGGGTACATGATGAGATTTGACAGCCCATCAGGGCCGGGTGG 1398
Db |||||||
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db |||||||
QY 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1458
Db |||||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db |||||||
QY 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
Db |||||||
QY 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 480
Db |||||||
QY 1519 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 1578
Db |||||||
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAATGGAGGAGAA 540
Db |||||||
QY 1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAATGGAGGAGAA 1638
Db |||||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
Db |||||||
QY 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1698
Db |||||||
QY 601 AGAAGATCTAGAACACAGAAACAAGTCAAGGTCAATTTCTCTCACTCATGTTGGTGGTAGT 660
Db |||||||
QY 1699 AGAAGATCTAGAACACAGAAACAAGTCAAGGTCAATTTCTCTCACTCATGTTGGTGGTAGT 1758
Db |||||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAACAATTAAGGTATTGGG 720
Db |||||||
QY 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAACAATTAAGGTATTGGG 1818
Db |||||||
QY 721 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAAGACAG 780
Db |||||||
QY 1819 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAAGACAT 1878
Db |||||||
QY 781 TTCTGACCAAGTGGAAAGCTCTGACCTTTCTCTGAGGAACTTCTGTTGTTGGCTACAGCT 840
Db |||||||
QY 1879 CCTTCTCAATGGCAACGCTTACTGAGAAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
Db |||||||
QY 841 GAAAGATGATGAATTAAGCCGGCAGGACCTTATTTGGAGGCGACTTTTCCAGCAGTTTCAGAA 900
Db |||||||
QY 1939 AAAAGAAAGATGCAGTGAACAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
Db |||||||
QY 901 GCAGAACGATGTACATAGGGCC--TTCAAGAGGGAATTTGAAAGAACTAAAGAACCTGTAAATC 958
Db |||||||
QY 1999 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 2058
Db |||||||

959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCCTTTTGA 1008
2059 GGGCAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
1009 AGGACTAGAGAAAATCTACAGAGGCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAA 1068
2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAA 2178
1069 TGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGTCAATAC----- 1110
2179 ACTTGAAAAGAGTACAGCACAGATTTACAGGCTGTCAACCCACTCAGCCATCACTAAC 2238
1111 ----- 1110
2239 ACAGACAACGTAAATGGAACACAGTAACACTACGGTGACCACAAGGGAACAGATCTCTGTAA 2298
1111 ---TGAGTGGGAAAAATTTGAACCTGCACTCCCGTCACTGGCAGAGAAAAATAGATGAGAC 1167
2299 GCATGCTCAAGAGGAACTTCCACCAACACCTCCCAAAAGAGAGGAGGAGTACTGTGA 2358
1168 CTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCA 1227
2359 TCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCA 2418
1228 AGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGGATCTCTCTATTGACTCTCTCA 1287
2419 AGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGGATCTCTCTATTGACTCTCTCA 2478
1288 AGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCCTCTCTGAAAGAGAACGT 1347
2479 AGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCCTCTCTGAAAGAGAACGT 2538
1348 GAGCCACGTCAATGACCTTGTGCGCCAGCTTACACCTTTGGGCACTCAGCTCTCACCGTA 1407
2539 GAGCCACGTCAATGACCTTGTGCGCCAGCTTACACCTTTGGGCACTCAGCTCTCACCGTA 2598
1408 TAACCTCAGCACTCTGGAAGACCTGGAACACCAAGATGGAAGCTTCTGAGGTGGCCGCGA 1467
2599 TAACCTCAGCACTCTGGAAGACCTGGAACACCAAGATGGAAGCTTCTGAGGTGGCCGCGA 2658
1468 GGACCGAGTCAGGCAGCTGCAATGAAGCCCCACAGGCACTTTGGTCCAGCATCTCAGCACTT 1527
2659 GGACCGAGTCAGGCAGCTGCAATGAAGCCCCACAGGCACTTTGGTCCAGCATCTCAGCACTT 2718
1528 TCTTTCCACGTCTGTCCAGGTCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTA 1587
2719 TCTTTCCACGTCTGTCCAGGTCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTA 2778
1588 CTATATCAACACGAGACTCAAAACAACTTGTGCGGACCATCCCAAAATGACAGAGCTCTA 1647
2779 CTATATCAACACGAGACTCAAAACAACTTGTGCGGACCATCCCAAAATGACAGAGCTCTA 2838
1648 CCAGTCTTTAGCTGACCTGAAATATGTCAAGATTTCAAGCTTATAGGACTGCCATGAAACT 1707
2839 CCAGTCTTTAGCTGACCTGAAATATGTCAAGATTTCAAGCTTATAGGACTGCCATGAAACT 2898
1708 CCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGGACCTGTGAGCTGCATGTATGC 1767
2899 CCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGGACCTGTGAGCTGCATGTATGC 2958
1768 CTTGGACCAAGCAAACTCAAGCAAAATGACCAAGCCCATGGATATCTGTCAGATTATTA 1827
2959 CTTGGACCAAGCAAACTCAAGCAAAATGACCAAGCCCATGGATATCTGTCAGATTATTA 3018
1828 TTGTTTGACCACTATTATGACCGCCCTGGAGCAAGAGCAAACTTTGGTCAACGTCCC 1887
3019 TTGTTTGACCACTATTATGACCGCCCTGGAGCAAGAGCAAACTTTGGTCAACGTCCC 3078
1888 TCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGG 1947
3079 TCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGG 3138

QY 1948 GAGGATCCGTGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGA 2007
Db 3139 GAGGATCCGTGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGA 3198
QY 2008 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCG 2067
Db 3199 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCG 3258
QY 2068 CAGGCTGGCCCTCTCTCTGTCATGATTTCTATCCAA 2101
Db 3259 CAGGCTGGCCCTCTCTCTGTCATGATTTCTATCCAA 3292

RESULT 14

US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match 76.2%; Score 1602; DB 10; Length 3999;
Best local similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
Db 900 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTCGAGTTCATTCATGAGAGAG 120
Db 960 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTCGAGTTCATTCATGAGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTTATCGTGGCTTCTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTTATCGTGGCTTCTTC 1079
QY 181 TGCTGAGGACACATTCGAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 240
Db 1080 TGCTGAGGACACATTCGAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1199
QY 301 TAATATTCTACAATTTGGGAAGTAAGTCTCTAAATTCAGATGGGAATTTATCAGAGATGAAGA 360
Db 1200 TAATATTCTACAATTTGGGAAGTAAGTCTCTAAATTCAGATGGGAATTTATCAGAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATTTATCAGAGATGAAGA 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATTTATCAGAGATGAAGA 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 1439

541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
1440 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCAATGATGTTGGTAGT 660
1500 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCAATGATGTTGGTAGT 1559
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGA--- 777
1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACCA 1679
778 ----- 777
1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 1739
778 ----- 777
1740 GGTGACACAACCTGTGGTTACTAAGGAACCTGCCATCTCCAAACTAGAAATGCCATCTTC 1799
778 ----- 777
1800 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAA 1859
778 ----- 777
1860 GTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCG 1919
778 ----- 777
1920 TAAGGAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCAAGA 1979
778 ----- 777
1980 CCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCA 2039
778 ----- 777
2040 AAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGA 2099
778 ----- 777
2100 TAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATT 2159
778 ----- CAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 831
2160 GGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 2219
832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 891
2220 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 2279
892 AGTTCAGAAGCAGACGATGTACATAGGCTTCAAGAGGGAATTGAAACTAAAGAAC 951
2280 AGTTCAGAAGCAGACGATGTACATAGGCTTCAAGAGGGAATTGAAACTAAAGAAC 2339
952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGG 1011
2340 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGG 2399
1012 ACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGT 1071
2400 ACTAGAGAACTCTACAGGAGGCCAGAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGT 2459
1072 CACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCT 1131
2460 CACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCT 2519

1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCA 1191
2520 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCA 2579
1192 AGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTG 1251
2580 AGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTG 2639
1252 GCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCTCAAGATCACCTCGAGAAAGTCAAGGC 1311
2640 GCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCTCAAGATCACCTCGAGAAAGTCAAGGC 2699
1312 ACTTCGAGAGAAAAATTTGGCCTCTCTGAAAGAGAACCTGAGCCACCTGCAATGACCTTCTCG 1371
2700 ACTTCGAGAGAAAAATTTGGCCTCTCTGAAAGAGAACCTGAGCCACCTGCAATGACCTTCTCG 2759
1372 CCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 1431
2760 CCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 2819
1432 GAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGA 1491
2820 GAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGA 2879
1492 AGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCCAGGGTCC 1551
2880 AGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCCAGGGTCC 2939
1552 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAAC 1611
2940 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAAC 2999
1612 AACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGCCTGAATAA 1671
3000 AACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGCCTGAATAA 3059
1672 TGTGAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCAGAGGCGCTTTG 1731
3060 TGTGAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCAGAGGCGCTTTG 3119
1732 CTTGGATCTTGTAGCCTGTGAGCTGATGATGCTGAGTCCCTTGACCAACACCTCAAGCA 1791
3120 CTTGGATCTTGTAGCCTGTGAGCTGATGATGCTGAGTCCCTTGACCAACACCTCAAGCA 3179
1792 AAATGACCAAGCCCATGGATATCTGAGATTAATTAATTTGTTGACCACTATTTATGACCG 1851
3180 AAATGACCAAGCCCATGGATATCTGAGATTAATTAATTTGTTGACCACTATTTATGACCG 3239
1852 CCTGGAGCAAGAGCACAACAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTGCTGAA 1911
3240 CCTGGAGCAAGAGCACAACAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTGCTGAA 3299
1912 CTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTTTTAA 1971
3300 CTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTTTTAA 3359
1972 AACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAA 2031
3360 AACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAA 3419
2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGGCGAGGCTGGGCTCCTTCTGCAATGA 2091
3420 GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGGCGAGGCTGGGCTCCTTCTGCAATGA 3479
2092 TTCTATCCAA 2101
3480 TTCTATCCAA 3489

RESULT 15
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1

GENERAL INFORMATION:

APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 4966
TYPE: DNA
ORGANISM: Homo sapiens
:-09-845-416-28

Query Match 76.2%; Score 1602; DB 10; Length 4966;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

1 GAGCTATGCTACACACAGGCTGTTATGTCAACCCTCTGACCCCTACACGAGCCCAT 60
|||||
1657 GAGCTATGCTACACACAGGCTGTTATGTCAACCCTCTGACCCCTACACGAGCCCAT 1716
|||||

61 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
|||||
1717 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1776
|||||

121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
|||||
1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1836
|||||

181 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
1837 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
|||||

241 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 300
|||||
1897 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 1956
|||||

301 TAATATTCTACAATTGGGAAGTAAGCTGATTTGAACAGGAAAAATTATCAGAAGATGAAGA 360
|||||
1957 TAATATTCTACAATTGGGAAGTAAGCTGATTTGAACAGGAAAAATTATCAGAAGATGAAGA 2016
|||||

361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
|||||

421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAACT 480
|||||
2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAACT 2136
|||||

481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAAGAAATGGAGGAAGA 540
|||||
2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 2196
|||||

541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
|||||
2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256
|||||

601 AGAAGATCTAGAACAAAGCAAGTCAGGCTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
|||||
2257 AGAAGATCTAGAACAAAGCAAGTCAGGCTCAATTTCTCTCACTCACATGGTGGTGGTAGT 2316
|||||

661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 720
|||||
2317 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 2376
|||||

721 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGTCTTTTACAAGA --- 777
|||||
2377 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGTCTTTTACAAGACCA 2436
|||||

QY 778 ----- 777
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACACTATTGGAGCCTCTCTACTCAGACTGTACTCT 2496
QY 778 ----- 777
Db 2497 GGTGACACACCTGTGGTTACTAAGGAAACTGCCATCTCAAAACTAGAAAATGCCATCTTC 2556
QY 778 ----- 777
Db 2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACACAGTTCCCTCCCTGGACCTGGAAAA 2616
QY 778 ----- 777
Db 2617 GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 2676
QY 778 ----- 777
Db 2677 TAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 2736
QY 778 ----- 777
Db 2737 CCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACACACCTGGATGAAACACAGCCA 2796
QY 778 ----- 777
Db 2797 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTTACAAAGACGTTTGG 2856
QY 778 ----- 777
Db 2857 TAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 2916
QY 778 ----- 831
Db 2917 GGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTG 2976
QY 832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGC 891
Db 2977 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGC 3036
QY 892 AGTTCAGAAGCAGAACCGATGTACATAGGCTTCAAGAGGGAAATTGAAAACTAAAGAAC 951
Db 3037 AGTTCAGAAGCAGAACCGATGTACATAGGCTTCAAGAGGGAAATTGAAAACTAAAGAAC 3096
QY 952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAG 1011
Db 3097 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAG 3156
QY 1012 ACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGGCCAGAAATGT 1071
Db 3157 ACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGGCCAGAAATGT 3216
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGCTCAATACTGAGTGGGAAAAAATTGAACCT 1131
Db 3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGCTCAATACTGAGTGGGAAAAAATTGAACCT 3276
QY 1132 GCCTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCA 1191
Db 3277 GCCTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCA 3336
QY 1192 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTG 1251
Db 3337 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTG 3396
QY 1252 GCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAAAGTCAAGGC 1311
Db 3397 GCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAAAGTCAAGGC 3456
QY 1312 ACITCAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 1371
Db 3457 ACITCAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 3516
QY 1372 CCAGCTTACCACCTTTGGGCATTCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGACCT 1431

```
b      ||||| 3517 CCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 3576
y      ||||| 1432 GAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGA 1491
b      ||||| 3577 GAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGA 3636
y      ||||| 1492 AGCCACACAGGGACTTTGGTCCAGCACTCTCAGCACTTCTTCCACGTCTGTCCAGGGTCC 1551
b      ||||| 3637 AGCCACACAGGGACTTTGGTCCAGCACTCTCAGCACTTCTTCCACGTCTGTCCAGGGTCC 3696
y      ||||| 1552 CTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCCACGAGACTCAAAC 1611
b      ||||| 3697 CTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCCACGAGACTCAAAC 3756
y      ||||| 1612 AACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 1671
b      ||||| 3757 AACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3816
y      ||||| 1672 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTG 1731
b      ||||| 3817 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTG 3876
y      ||||| 1732 CTGGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCA 1791
b      ||||| 3877 CTGGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCA 3936
y      ||||| 1792 AAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACCTATTTATGACCG 1851
b      ||||| 3937 AAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACCTATTTATGACCG 3996
y      ||||| 1852 CCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGGTGGATATGTGTCTGAA 1911
b      ||||| 3997 CCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGGTGGATATGTGTCTGAA 4056
y      ||||| 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAA 1971
b      ||||| 4057 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAA 4116
y      ||||| 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 2031
b      ||||| 4117 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 4176
y      ||||| 2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCATGA 2091
b      ||||| 4177 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCATGA 4236
y      ||||| 2092 TTCTATCCAA 2101
b      ||||| 4237 TTCTATCCAA 4246
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 nucleic - nucleic search, using sw model

in on: April 4, 2004, 11:55:32 ; Search time 542.141 Seconds
(without alignments)
16463.377 Million cell updates/sec

itle: US-09-845-416-12_COPY_900_3000
ffect score: 2101
quence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

oring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

arched: 3373863 seqs, 2124099041 residues

tal number of hits satisfying chosen parameters: 6747726

imum DB seq length: 0
imum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : N_Geneseq_29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002s: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB	ID	Description
1	2101	100.0	3510	6	AAD37240	Aad37240 Human dys
2	2101	100.0	4476	6	AAD37259	Aad37259 Adeno-ass
3	1760.4	83.8	5339	6	ABK81998	Abk81998 DNA encod
4	1750.8	83.3	3531	6	AAD37238	Aad37238 Human dys
5	1750.8	83.3	4498	6	AAD37258	Aad37258 Adeno-ass
6	1743	83.0	3858	6	AAD37237	Aad37237 Human dys
7	1743	83.0	4825	6	AAD37257	Aad37257 Adeno-ass
8	1743	83.0	4848	6	AAD37263	Aad37263 Adeno-ass
9	1743	83.0	5060	6	AAD37264	Aad37264 Adeno-ass
10	1641	78.1	4414	6	AAD37260	Aad37260 Adeno-ass
11	1629	77.5	3446	6	AAD37242	Aad37242 Human dys
12	1629	77.5	5462	6	ABK81999	Abk81999 DNA encod
13	1611.6	76.7	5417	6	ABK81997	Abk81997 DNA encod
14	1602	76.2	3999	6	AAD37234	Aad37234 Human dys
15	1602	76.2	4966	6	AAD37256	Aad37256 Adeno-ass
16	1602	76.2	4990	6	AAD37262	Aad37262 Adeno-ass
17	1419	67.5	4182	6	AAD37230	Aad37230 Human dys
18	1419	67.5	5149	6	AAD37255	Aad37255 Adeno-ass
19	1328.8	63.2	2169	6	AAD37232	Aad37232 Human dys
20	1328.8	63.2	5952	5	AAD06794	Aad06794 Human dys
21	1328.8	63.2	8689	6	ABK82000	Abk82000 DNA encod
22	1328.8	63.2	11058	6	AAD37229	Aad37229 Human dys
23	1328.8	63.2	11241	6	ABK82005	Abk82005 cDNA enco

24	1328.8	63.2	11443	6	ABK82002	Abk82002 DNA encod
25	1328.8	63.2	12923	1	AAN90338	Aan90338 Sequence
26	1328.8	63.2	13957	6	ABK81959	Abk81959 cDNA enco
27	1328.8	63.2	13957	6	ABT10904	Abt10904 Human bre
28	1328.8	63.2	13957	6	ABN95786	Abn95786 Gene #228
29	1328.8	63.2	13957	6	ABS69900	Abs69900 Human dys
30	1324	63.0	1821	6	AAD37241	Aad37241 Human dys
31	1317.8	62.7	13977	6	ABS70403	Abs70403 Human bon
32	1317	62.7	4402	3	AAD37243	Aad37243 Human dys
33	1196.6	57.0	4402	3	AAD37235	Aad37235 Human dys
34	1149.6	54.7	13815	6	ABK81960	Abk81960 cDNA enco
35	1149.6	54.7	13815	6	ABI99799	Abi99799 Mouse isc
36	1149.6	54.7	19307	2	AAT27558	Aat27558 Shuttle v
37	1148	54.6	13815	2	AAV18885	Aav18885 Mus muscu
38	937	44.6	1434	6	AAD37243	Aad37243 Human dys
39	785.4	37.4	1991	6	AAD37231	Aad37231 Human dys
40	777	37.0	1667	6	AAD37235	Aad37235 Human dys
41	718.8	34.2	4075	3	AAD37235	Aad37235 Human dys
42	667.2	31.8	3747	3	AAD37235	Aad37235 Human dys
43	618.6	29.4	3275	1	AAN97129	Aan97129 Partial s
44	613.6	29.2	3163	3	AAD37235	Aad37235 Human dys
45	555.6	26.4	10705	7	ABT41896	Abt41896 Toxicity

ALIGNMENTS

RESULT 1
AAD37240
ID AAD37240 standard; DNA; 3510 BP.
XX
AC AAD37240;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
(XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 51-52; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CR domain) and 11047-11058 (dystrophin last 3 amino acids)
Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
Query Match 100.0%; Score 2101; DB 6; Length 3510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACAGGAGCCCAT 60
900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACAGGAGCCCAT 959
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1019
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 180
1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 1079
181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1080 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
1140 CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1199
301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
1200 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1259
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCCCTCAGGGTAGC 420
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAAGATGGGAATGCCCTCAGGGTAGC 1319
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 540
1380 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA 1439
541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
1440 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCACTCACTCAATGGTGGTGTAGT 660
1500 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCACTCACTCAATGGTGGTGTAGT 1559
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 780
1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 1679
781 TTCTGACAGTGGGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCT 840
1680 TTCTGACAGTGGGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCT 1739
841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGGACTTTCCAGCAGTTTCAGAA 900
1740 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGGACTTTCCAGCAGTTTCAGAA 1799
901 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAATCAT 960

1800 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAACTGTAATCAT 1859
961 GAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGACTAGAGAA 1020
1860 GAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGACTAGAGAA 1919
1021 ACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1080
1920 ACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1979
1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTTGAACCTGCACCTCCGC 1140
1980 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTTGAACCTGCACCTCCGC 2039
1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1200
2040 TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 2099
1201 GGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGTGATCAAGGGATCTTGGCAGCCCCGT 1260
2100 GGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGTGATCAAGGGATCTTGGCAGCCCCGT 2159
1261 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1320
2160 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2219
1321 AGAAATTCGGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1380
2220 AGAAATTCGGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 2279
1381 CACTTTTGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAG 1440
2280 CACTTTTGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAG 2339
1441 ATGGAAGCTTCTGCAGGTGGCGTGCAGGACCAGTCCAGGCTGCTGATGAAGCCCCACAG 1500
2340 ATGGAAGCTTCTGCAGGTGGCGTGCAGGACCAGTCCAGGCTGCTGATGAAGCCCCACAG 2399
1501 GGACTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTCTGTCCAGGGTCCCTTGGGAGAG 1560
2400 GGACTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTCTGTCCAGGGTCCCTTGGGAGAG 2459
1561 AGCCATCTGCGCCAAAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGCTG 1620
2460 AGCCATCTGCGCCAAAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGCTG 2519
1621 GGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATATGTCAGATT 1680
2520 GGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATATGTCAGATT 2579
1681 CTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCT 1740
2580 CTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAAGAGGCCCTTTTGTGGATCT 2639
1741 CTTGAGCCTGTCAAGTGCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCA 1800
2640 CTTGAGCCTGTCAAGTGCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCA 2699
1801 GCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 1860
2700 GCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 2759
1861 AGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCTGGGATATGTGTGAACCTGGCTGCT 1920
2760 AGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCTGGGATATGTGTGAACCTGGCTGCT 2819
1921 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTGCTTTTAAACTGGCAT 1980
2820 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTGCTTTTAAACTGGCAT 2879
1981 CATTTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
2880 CATTTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2939

Y 2041 AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGCCCTCCTTCTGCATGATTTCTATCCA 2100
b 2940 AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGCCCTCCTTCTGCATGATTTCTATCCA 2999
Y 2101 A 2101
b 3000 A 3000

ESULT 2
AD37259
D AAD37259 standard; DNA; 4476 BP.
X
C AAD37259;
X
T 21-AUG-2002 (first entry)
X
E Adeno-associated virus vector plasmid, AAV-MCK-3510.
X
X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X
S Homo sapiens.
S Unidentified.
S Chimeric.

X WO200183695-A2.
X
N 08-NOV-2001.
D
X 27-APR-2001; 2001WO-US013677.
F
X 28-APR-2000; 2000US-020077P.
R
X (XIAO/) XIAO X.
A
X Xiao X;
I
X WPI; 2002-049342/06.
R
X

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 63-65; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

Query Match 100.0%; Score 2101; DB 6; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGGAGCCCAT 60
b 1656 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGGAGCCCAT 1715
Y 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120

Db 1716 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1775
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
Db 1776 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1835
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1836 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1895
QY 241 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTGG 300
Db 1896 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTGG 1955
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 360
Db 1956 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 2015
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 2016 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 2075
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 480
Db 2076 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 2135
QY 481 GAAAGAGTTGAATGACTGGTAAACAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 540
Db 2136 GAAAGAGTTGAATGACTGGTAAACAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 2195
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACTAAGGTGCTTCA 600
Db 2196 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACTAAGGTGCTTCA 2255
QY 601 AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db 2256 AGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 2315
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACCTTAAGGTATTGGG 720
Db 2316 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACCTTAAGGTATTGGG 2375
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 780
Db 2376 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 2435
QY 781 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGTTGTTGGCTACAGCT 840
Db 2436 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGTTGTTGGCTACAGCT 2495
QY 841 GAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA 900
Db 2496 GAAAGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA 2555
QY 901 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTAAAGAACCTGTATCAT 960
Db 2556 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTAAAGAACCTGTATCAT 2615
QY 961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGACGACCTTTTGGAGGACTAGAGAA 1020
Db 2616 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGACGACCTTTTGGAGGACTAGAGAA 2675
QY 1021 ACTCTACAGGAGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGT 1080
Db 2676 ACTCTACAGGAGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGT 2735
QY 1081 TCTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGAAAAAATTAAGAACTGCACTCCGC 1140
Db 2736 TCTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGAAAAAATTAAGAACTGCACTCCGC 2795
QY 1141 TGACTGGCAGAGAAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1200

2796 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGCTCCAGGAACCTTCAAGAGGCCAC 2855
1201 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 1260
2856 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 2915
1261 GGGCGATCTCCTCATTCATTCCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
2916 GGGCGATCTCCTCATTCATTCCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2975
1321 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTAC 1380
2976 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTAC 3035
1381 CACTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
3036 CACTTTGGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 3095
1441 ATGGAAGCTTCTGCAGGTGCGCGTGCAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAG 1500
3096 ATGGAAGCTTCTGCAGGTGCGCGTGCAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAG 3155
1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAG 1560
3156 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAG 3215
1561 AGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCCAGAGACTCAACAACTTGCTG 1620
3216 AGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCCAGAGACTCAACAACTTGCTG 3275
1621 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATTAATGTCAGATT 1680
3276 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATTAATGTCAGATT 3335
1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTGCTTGGATCT 1740
3336 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTGCTTGGATCT 3395
1741 CTTGAGCCTGTGAGCTGCTGATGCTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCA 1800
3396 CTTGAGCCTGTGAGCTGCTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCA 3455
1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 1860
3456 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 3515
1861 AGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGATATGCTGTAAGTGGCTGCT 1920
3516 AGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGATATGCTGTAAGTGGCTGCT 3575
1921 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTCCTGCTCTTTTAAACTGGCAT 1980
3576 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTCCTGCTCTTTTAAACTGGCAT 3635
1981 CATTTCCTGTGTAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
3636 CATTTCCTGTGTAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 3695
2041 AAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCA 2100
3696 AAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCA 3755
2101 A 2101
3756 A 3756

RESULT 3
ABK81998
ID ABK81998 standard; DNA; 5339 BP.
AC
XX ABK81998;
XX

DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Example 6; Fig 13; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;
Query Match 83.8%; Score 1760.4; DB 6; Length 5339;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 191; Indels 15; Gaps 2;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGTATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGTATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGGTGG 300
Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGGTGG 1398
QY 301 TAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATATATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATATATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGAGTGGAAATGCCTCAGGGTAGC 420

1459 AACTGAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
1519 TAGCATGGAACAAACAAAGCAATTTACATCATAGATTACTGCAACAGTTCCCTCCCTGGACCT 1578
481 GAAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAAAGAACACAGGAAAATGGAGGAAGA 540
1579 GGAAGAGTTTCTTGCTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGC 1638
541 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGTGTCTTCA 600
1639 TACCCGTAAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATG 1698
601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 660
1699 GCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACTGGATGAAA 1758
661 TGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACT 708
1759 CAGCCAAAATACTCTGAGATCCCTGGAAGTTCGATGATGAGTCTCTTTTACAAAGAGC 1818
709 TAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGTGGGT 765
1819 TTTGGATAACATGAACITCAAGTGGAGTGAATTCGGAAGAGTCTCTCAACATTAGGTC 1878
766 TCTTTTACAAGACAGTTCTGACCAGTGGAGCGTCTGCAACCTTCTCTGACAGGAACCTTCT 825
1879 CCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCAACCTTCTCTGACAGGAACCTTCT 1938
826 GGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCACTATTGGAGCGGACTT 885
1939 GGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCACTATTGGAGCGGACTT 1998
886 TCCAGCAGTTTCAAGAGCAGAACGATCTACATAGGGCCCTTCAAGAGGGAATTTGAAAACATAA 945
1999 TCCAGCAGTTTCAAGAGCAGAACGATCTACATAGGGCCCTTCAAGAGGGAATTTGAAAACATAA 2058
946 AGAACCTGTAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 1005
2059 AGAACCTGTAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 2118
1006 GGAAGGACTAGAGAACTCTACAGGAGCCCGAGAGAGCTGCTCCTGAGGAGAGAGCCCA 1065
2119 GGAAGGACTAGAGAACTCTACAGGAGCCCGAGAGAGCTGCTCCTGAGGAGAGAGCCCA 2178
1066 GAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAATTT 1125
2179 GAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAATTT 2238
1126 GAACCTGCACCTCGGCTGACTGGCAGAGAAAATATAGATGAGACCTTTGAAAGACTCCAGGA 1185
2239 GAACCTGCACCTCGGCTGACTGGCAGAGAAAATATAGATGAGACCTTTGAAAGACTCCAGGA 2298
1186 ACTTCAAGAGGGCCACGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGG 1245
2299 ACTTCAAGAGGGCCACGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGG 2358
1246 ATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCTCAAGATCACCTCGAGAAAGT 1305
2359 ATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCTCAAGATCACCTCGAGAAAGT 2418
1306 CAAGGCACCTTCGAGGAGAAATTTGGCCCTTGAAAGAGAACCTGAGCCACCTCAATGACCT 1365
2419 CAAGGCACCTTCGAGGAGAAATTTGGCCCTTGAAAGAGAACCTGAGCCACCTCAATGACCT 2478
1366 TGCTGCCAGCTTACCACCTTTGGGCATTTAGCTCTCACCCGTATTAACCTCAGCACTCTGGA 1425
2479 TGCTGCCAGCTTACCACCTTTGGGCATTTAGCTCTCACCCGTATTAACCTCAGCACTCTGGA 2538
1426 AGACCTGAACACCCAGATGGAAGCTTCTGAGGTTGGCCGTGAGGACCGAGTCAGGCAGCT 1485
2539 AGACCTGAACACCCAGATGGAAGCTTCTGAGGTTGGCCGTGAGGACCGAGTCAGGCAGCT 2598

QY 1486 GCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTGTGTCCA 1545
Db 2599 GCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTGTGTCCA 2658
QY 1546 GGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGAC 1605
Db 2659 GGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGAC 2718
QY 1606 TCAAAACAACCTTGTCTGGGACCATCCCAAAAATGACAGAGCTTACAGTCTTTAGCTGACCT 1665
Db 2719 TCAAAACAACCTTGTCTGGGACCATCCCAAAAATGACAGAGCTTACAGTCTTTAGCTGACCT 2778
QY 1666 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGC 1725
Db 2779 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGC 2838
QY 1726 CCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCTTGGACCAGCAACCT 1785
Db 2839 CCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCTTGGACCAGCAACCT 2898
QY 1786 CAAGCAAAATGACAGCCCATGGATATCTGAGATTATTAATTTTGACCACTATTTA 1845
Db 2899 CAAGCAAAATGACAGCCCATGGATATCTGAGATTATTAATTTTGACCACTATTTA 2958
QY 1846 TGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGCTCCTCTCTGCGTGGATATGTG 1905
Db 2959 TGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGCTCCTCTCTGCGTGGATATGTG 3018
QY 1906 TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGCTGCTGTC 1965
Db 3019 TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGCTGCTGTC 3078
QY 1966 TTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCT 2025
Db 3079 TTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCT 3138
QY 2026 TTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCT 2085
Db 3139 TTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCT 3198
QY 2086 GCATGATTCTATCCAA 2101
Db 3199 GCATGATTCTATCCAA 3214

RESULT 4
AAD37238

ID AAD37238 standard; DNA; 3531 BP.

XX AAD37238;

AC AAD37238;

XX 21-AUG-2002 (first entry)

DT Human dystrophin minigene delta3531.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

OS Homo sapiens.

XX WO200183695-A2.

PN 08-NOV-2001.

PD 27-APR-2001; 2001WO-US013677.

XX 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

PI Xiao X;

WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
Example 1; Page 50-51; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus, hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;
Query Match 83.3%; Score 1750.8; DB 6; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGGAGCCCAATT 60
900 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGGAGCCCAATT 959
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
121 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTTC 180
1020 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTTC 1079
181 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1080 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAGTTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 300
1140 CCAGTTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 1199
301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
1200 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1259
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAAATCA 474
1320 TAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCAACAGTTCCCCCT 1379
475 GAAACTGAAGAGATTGAATGACTGGCTTAACAAAAACAGAGAAAGAAACAAAGGAAAAATGGA 534
1380 GGACCTGGAAAAAGTTTCTTGGCTGGTTACAGAAGCTGAAACAACTGCCAATGTCCTTACA 1439
535 GGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGT 594
1440 GGATGCTACCCGTAAGGAAAGGCTCTTAGAGACTCCAAAGGAGTAAAGAGCTGATGAA 1499
595 GCTTCAAGAAGATCTAGAACAAAGACAGTCAAGGTCATTTCTCACTCAGATGGTGGT 654
1500 ACAATGGCAAGACCTCCAGGGTGAATTGAAGCTCACACAGATGTTTATCACAAACCTTGA 1559

QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGA 702
Db 1560 TGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACA 1619
QY 703 ACAACTTAAGG--TATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCG 759
Db 1620 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACAT 1679
QY 760 CTGGGTTCTTTTACAAGACAGTCTTGACCCAGTGGGAAGCGTCTGCACCTTTCTCTGCAGGA 819
Db 1680 TAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTGGGAAGCGTCTGCACCTTTCTCTGCAGGA 1739
QY 820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 879
Db 1740 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 1799
QY 880 CGACTTTCAGCAGTTTCAGAACGAGAACGATGTATACATAGGGCCTTCAAGAGGGAATTGAA 939
Db 1800 CGACTTTCAGCAGTTTCAGAACGAGAACGATGTATACATAGGGCCTTCAAGAGGGAATTGAA 1859
QY 940 AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 999
Db 1860 AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
QY 1000 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCTCTGAGGAGAG 1059
Db 1920 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCTCTGAGGAGAG 1979
QY 1060 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 1119
Db 1980 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2039
QY 1120 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCCCTTGAAGACT 1179
Db 2040 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCCCTTGAAGACT 2099
QY 1180 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 1239
Db 2100 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2159
QY 1240 CAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 1299
Db 2160 CAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2219
QY 1300 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAC 1359
Db 2220 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAC 2279
QY 1360 TGACCTTGTGCGCCAGCTTACCCTTTTGGGCAATTCAGCTCTCAACGTTATACCTCAGCAC 1419
Db 2280 TGACCTTGTGCGCCAGCTTACCCTTTTGGGCAATTCAGCTCTCAACGTTATACCTCAGCAC 2339
QY 1420 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 1479
Db 2340 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 2399
QY 1480 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1539
Db 2400 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2459
QY 1540 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 1599
Db 2460 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 2519
QY 1600 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 1659
Db 2520 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 2579
QY 1660 TGACCTGAATATGTGATGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
Db 2580 TGACCTGAATATGTGATGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639

1720 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 1779
|||||
2640 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 2699
|||||
1780 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCCTGCAGATTATTAAATGTTTGACCCAC 1839
|||||
2700 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCCTGCAGATTATTAAATGTTTGACCCAC 2759
|||||
1840 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1899
|||||
2760 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 2819
|||||
1900 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1959
|||||
2820 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 2879
|||||
1960 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2019
|||||
2880 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2939
|||||
2020 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 2079
|||||
2940 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 2999
|||||
2080 CCTTCTGTCATGATTTCTATCCAA 2101
|||||
3000 CCTTCTGTCATGATTTCTATCCAA 3021
|||||

RESULT 5

AD37258
AAD37258 standard; DNA; 4498 BP.

AAD37258;

21-AUG-2002 (first entry)

Adeno-associated virus vector plasmid, AAV-MCK-3531.

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.

Homo sapiens.
Unidentified.
Chimeric.

WO200183695-A2.

08-NOV-2001.

27-APR-2001; 2001WO-US013677.

28-APR-2000; 2000US-0200777P.

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.

Example 1; Page 62-63; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)

comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;

Query Match 83.3%; Score 1750.8; DB 6; Length 4498;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
|||||
1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1716
|||||
61 TCCTTCACAGCATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
|||||
1717 TCCTTCACAGCATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
|||||
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTCTCTTC 180
|||||
1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTCTCTTC 1836
|||||
181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
|||||
241 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
1897 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
|||||
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATATCAGAAGATGAAGA 360
|||||
1957 TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATATCAGAAGATGAAGA 2016
|||||
361 AACTGAAGTACAAGACAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
2017 AACTGAAGTACAAGACAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
|||||
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAATCA 474
|||||
2077 TAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCAACAGTTTCCCCCT 2136
|||||
475 GAAACTGAAAGAGTTGAATGACTGGCTTACAAAAACAGAAAGAAAGAAATGGA 534
|||||
2137 GGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCAATGTCTCTACA 2196
|||||
535 GGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGT 594
|||||
2197 GGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAAGAGCTGATGAA 2256
|||||
595 GCTTCAAGAAGATCTAGAACAAGACAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGT 654
|||||
2257 ACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGA 2316
|||||
655 GGTAG-----TTGATGAATCTAGTGGAGATCAGCGCACTGCTTTTGAAGA 702
|||||
2317 TGAACACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTACA 2376
|||||
703 ACAACTTAAGG---TATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCG 759
|||||
2377 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 2436
|||||
760 CTGGGTTCTTTTACAAGACAGTTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 819
|||||
2437 TAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 2496
|||||
820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACACCTATTGGAGG 879
|||||
2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACACCTATTGGAGG 2556
|||||

880 CGACTTCCAGCAGTTTCAAGACGAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 939
2557 CGACTTCCAGCAGTTTCAAGACGAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 2616
940 AACTAAAGAACCTGTATCATGAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCA 999
2617 AACTAAAGAACCTGTATCATGAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCA 2676
1000 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 1059
2677 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 2736
1060 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 1119
2737 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2796
1120 AAAATTGAACCTGCCTCGCTGACTGGCGAGAGAAATAGATGAGACCCCTTGAAGAGCT 1179
2797 AAAATTGAACCTGCCTCGCTGACTGGCGAGAGAAATAGATGAGACCCCTTGAAGAGCT 2856
1180 CCAGGAACCTTCAAGAGGCGACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 1239
2857 CCAGGAACCTTCAAGAGGCGACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2916
1240 CAAGGATCCTGGCAGCCGCTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 1299
2917 CAAGGATCCTGGCAGCCGCTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2976
1300 GAAAGTCAAGGCACCTTCAGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCAGTCAA 1359
2977 GAAAGTCAAGGCACCTTCAGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCAGTCAA 3036
1360 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATATACCTCAGCAC 1419
3037 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATATACCTCAGCAC 3096
1420 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCA 1479
3097 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCA 3156
1480 GCAGCTGCATGAAGCCACAGGGGACTTTGTGTCAGCATCTCAGACACTTTCTTCCACGTC 1539
3157 GCAGCTGCATGAAGCCACAGGGGACTTTGTGTCAGCATCTCAGACACTTTCTTCCACGTC 3216
1540 TGTCAGGGTCCCTGGGAGAGAGCCATCTGCGCCAAACAAAGTGCCTTACTATATCAACCA 1599
3217 TGTCAGGGTCCCTGGGAGAGAGCCATCTGCGCCAAACAAAGTGCCTTACTATATCAACCA 3276
1600 CGAGACTCAAAACAACTTGTGTCAGCATCTCAGACACTTCTACAGTCTTACAGTCTTTAGC 1659
3277 CGAGACTCAAAACAACTTGTGTCAGCATCTCAGACACTTCTACAGTCTTACAGTCTTTAGC 3336
1660 TGACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 1719
3337 TGACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 3396
1720 GAAGGCCCTTTGCTTGGATCTTGTGAGCCTGTGAGTGCATGTGATGCCCTTGGACACGCA 1779
3397 GAAGGCCCTTTGCTTGGATCTTGTGAGCCTGTGAGTGCATGTGATGCCCTTGGACACGCA 3456
1780 CAACCTCAAGCAAAATGACAGCCAGCCCATGATATCCTGAGATTAATTAATTTGACCCAC 1839
3457 CAACCTCAAGCAAAATGACAGCCAGCCCATGATATCCTGAGATTAATTAATTTGACCCAC 3516
1840 TATTATGACCGCTGGAGCAAGAGCACAATAATTTGGTCAACGTCCCTCTCTGCGTGA 1899
3517 TATTATGACCGCTGGAGCAAGAGCACAATAATTTGGTCAACGTCCCTCTCTGCGTGA 3576
1900 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGGAGGATCCGTGT 1959
3577 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGGAGGATCCGTGT 3636

1960 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 2019
3637 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 3696
2020 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGTTTGTGACCAAGCGAGGCTGGGCCT 2079
3697 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGTTTGTGACCAAGCGAGGCTGGGCCT 3756
2080 CCTTCTGCATGATTCTATCCAA 2101
3757 CCTTCTGCATGATTCTATCCAA 3778
RESULT 6
AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
WPI; 2002-049342/06.
XX
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.
XX
Example 1; Page 48-49; 71pp; English.
XX
The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta3849 containing nucleotides 1-1568 (N-terminus,
hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;
Query Match 83.0%; Score 1743; DB 6; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTATGTCACACCTCTGACCCCTACACGAGCCCAT 60
DB 900 GAGCTATGCCTACACACAGGCTGCTTATGTATGTCACACCTCTGACCCCTACACGAGCCCAT 959

Y 61 TCCTTACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
b 960 TCCTTACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1019
Y 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATATCGTGGCTCTTTC 180
b 1020 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATATCGTGGCTCTTTC 1079
Y 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
b 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 1139
Y 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
b 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1199
Y 301 TAAATATTCTACATTTGGGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 360
b 1200 TAAATATTCTACATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259
Y 361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
b 1260 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319
Y 421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAAT 480
b 1320 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAAT 1379
Y 481 GAAAGAGTTGAATGACTGGCTAACAAAACACAGAAGAAAGAAACAAGGAAAATGGAGGAAGA 540
b 1380 GAAAGAGTTGAATGACTGGCTAACAAAACACAGAAGAAAGAAACAAGGAAAATGGAGGAAGA 1439
Y 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACATAAGGTGCTTCA 600
b 1440 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACATAAGGTGCTTCA 1499
Y 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGGTAGT 660
b 1500 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGGTAGT 1559
Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAAGTATTGGG 720
b 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAAGTATTGGG 1619
Y 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGA --- 777
b 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGACAC 1679
Y 778 ----- 777
b 1680 TCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAAGTTTCTTGGCTGGCTTACAGA 1739
Y 778 ----- 777
b 1740 AGCTGAAACAACCTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1799
Y 778 ----- 777
b 1800 CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATGAAGC 1859
Y 778 ----- 777
b 1860 TCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAAATACTCTGAGATCCCTGGA 1919
Y 778 ----- 777
b 1920 AGGTTCCGATGATGAGTCCCTGTTTACAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1979
Y 778 ----- CAGTTCTGACCAAGTG 792
b 1980 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTG 2039

QY 793 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGTGAAAGATGATGA 852
Db 2040 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGTGAAAGATGATGA 2099
QY 853 ATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAACGAGAACGATGT 912
Db 2100 ATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAACGAGAACGATGT 2159
QY 913 ACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTATCATGACTCTTTGA 972
Db 2160 ACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTATCATGACTCTTTGA 2219
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGA 1032
Db 2220 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGA 2279
QY 1033 GCCACAGAGAGCTGCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
Db 2280 GCCACAGAGAGCTGCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2339
QY 1093 GGCTGAGGAGGTCAATFACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
Db 2340 GGCTGAGGAGGTCAATFACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG 2399
QY 1153 AAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1212
Db 2400 AAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2459
QY 1213 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 1272
Db 2460 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 2519
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 1332
Db 2520 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 2579
QY 1333 TCTGAAAGAGAACGTTGAGCCACGTTCAATGACCTTGTGCGCCAGCTTACCACCTTTGGGCA 1392
Db 2580 TCTGAAAGAGAACGTTGAGCCACGTTCAATGACCTTGTGCGCCAGCTTACCACCTTTGGGCA 2639
QY 1393 TCAGCTCTCACCCGTATAACCTCAGCACTCTTGAAGACCTGAAACACACAGATGGAAGCTTCT 1452
Db 2640 TCAGCTCTCACCCGTATAACCTCAGCACTCTTGAAGACCTGAAACACACAGATGGAAGCTTCT 2699
QY 1453 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGTGCATGAAGCCACACAGGGACTTTGGTCC 1512
Db 2700 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGTGCATGAAGCCACACAGGGACTTTGGTCC 2759
QY 1513 AGCATCTCAGCACTTTCTTCCACGTTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db 2760 AGCATCTCAGCACTTTCTTCCACGTTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2819
QY 1573 AAACAAAGTGCCCTACTATATCAACACAGAGACTCAACAACTTGTCTGGACCATCCCCAA 1632
Db 2820 AAACAAAGTGCCCTACTATATCAACACAGAGACTCAACAACTTGTCTGGACCATCCCCAA 2879
QY 1633 AATGACAGAGCTTACCAGTCTTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAG 1692
Db 2880 AATGACAGAGCTTACCAGTCTTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAG 2939
QY 1693 GACTGCCATGAACTCCGAAGACTGCAAGAGGCCCTTTTGTGGATCTCTTGAGCCTGTC 1752
Db 2940 GACTGCCATGAACTCCGAAGACTGCAAGAGGCCCTTTTGTGGATCTCTTGAGCCTGTC 2999
QY 1753 AGCTGCATGTGATGCTTGGACCAGCAACACCTCAAGCAAAAATGACCCAGCCCATGGATAT 1812
Db 3000 AGCTGCATGTGATGCTTGGACCAGCAACACCTCAAGCAAAAATGACCCAGCCCATGGATAT 3059
QY 1813 CCTGCAGATTATTAATGTTTGGACCACTATTTATGACCGCTGGAGCAAGAGCAACAA 1872
Db 3060 CCTGCAGATTATTAATGTTTGGACCACTATTTATGACCGCTGGAGCAAGAGCAACAA 3119
QY 1873 TTTGGTCAACGTCCTCTCTGCGTGGATGTGTCTGAACCTGGCTGCTGAATGTTTATGA 1932

db 3120 TTTGGTCAACGTCCTCTCTCGCTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGA 3179
dy 1933 TACGGGACGAACAGGGAGGATCCGTGTCCCTGTCTTTTAAAACTGGCATATTTCCCTGTG 1992
db 3180 TACGGGACGAACAGGGAGGATCCGTGTCCCTGTCTTTTAAAACTGGCATATTTCCCTGTG 3239
dy 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
db 3240 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 3299
dy 2053 ATTTTGTGACCGCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
db 3300 ATTTTGTGACCGCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 3348

RESULT 7

AAD37257 standard; DNA; 4825 BP.
AAD37257;
21-AUG-2002 (first entry)
Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
Homo sapiens.
Unidentified.
Chimeric.
WO200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.

Example 1; Page 61-62; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match 83.0%; Score 1743; DB 6; Length 4825;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTGGCAGTTTCATTGATGGAGAG 120
Db 1717 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTGGCAGTTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1897 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGGAAAAACAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGTTCCTTTTACAAGA --- 777
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGTTCCTTTTACAAGACAC 2436
QY 778 ----- 777
Db 2437 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA 2496
QY 778 ----- 777
Db 2497 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTPAGAAGA 2556
QY 778 ----- 777
Db 2557 CTCCAAGGGAGTAAAGAGCTGTATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616
QY 778 ----- 777
Db 2617 TCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGA 2676
QY 778 ----- 777
Db 2677 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2736

778 -----CAGTTCTGACCAGTG 792
b 2737 TGAAC TTCGGA AAAAGTCTCTCAACATAGGTCCCATTTTGAAGCCAGTTCTGACCAGTG 2796
y 793 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGA 852
b 2797 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGA 2856
y 853 ATTAAGCCGGCAGGACCTATTGGAGGCGGACTTTCCAGCGAGTTCCAGAACGACGACGATGT 912
b 2857 ATTAAGCCGGCAGGACCTATTGGAGGCGGACTTTCCAGCGAGTTCCAGAACGACGACGATGT 2916
y 913 ACATAGGGCCCTTCAAGAGGGGAATTGA AAAACTAAAGAACCTGTAAATCATGACTACTTTGA 972
b 2917 ACATAGGGCCCTTCAAGAGGGGAATTGA AAAACTAAAGAACCTGTAAATCATGACTACTTTGA 2976
y 973 GACTGTACGAATATTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACACAGGA 1032
b 2977 GACTGTACGAATATTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACACAGGA 3036
y 1033 GCCCAGAGAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
b 3037 GCCCAGAGAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 3096
y 1093 GGCTGAGGAGGTCAATATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
b 3097 GGCTGAGGAGGTCAATATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 3156
y 1153 AAAAAATAGATGAGACCTTGAAGGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1212
b 3157 AAAAAATAGATGAGACCTTGAAGGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 3216
y 1213 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1272
b 3217 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3276
y 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCC 1332
b 3277 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCC 3336
y 1333 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 1392
b 3337 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 3396
y 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 1452
b 3397 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 3456
y 1453 GCAGTGGCCGTCGAGGACCGAGTCAGGCACTGATGAAGCCACAGGCACTTTGGTCC 1512
b 3457 GCAGTGGCCGTCGAGGACCGAGTCAGGCACTGATGAAGCCACAGGCACTTTGGTCC 3516
y 1513 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGACCATCTCGCC 1572
b 3517 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGACCATCTCGCC 3576
y 1573 AAACAAAGTGCCTACTATATACCAACGAGACTCAAACTTGTCTGGGACCATCCCAA 1632
b 3577 AAACAAAGTGCCTACTATATACCAACGAGACTCAAACTTGTCTGGGACCATCCCAA 3636
y 1633 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692
b 3637 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3696
y 1693 GACTGCCATGAAACTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
b 3697 GACTGCCATGAAACTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3756
y 1753 AGCTGCATGTGATGCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCATGGATAT 1812
b 3757 AGCTGCATGTGATGCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCATGGATAT 3816
y 1813 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGGCTGGAGCAAGAGCAACAAC 1872

Db 3817 CCTGCAGATTATTAATTGTTTGGACCACTATTATGACCGCTGGAGCAAGAGCAACAAC 3876
Qy 1873 TTTGGTCAACGTCCTCTCTGGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 1932
Db 3877 TTTGGTCAACGTCCTCTCTGGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 3936
Qy 1933 TACGGACGAACAGGAGGATCCGTGTCCTCTCTTTTAAAACTGGCATCATTTCCCTGTG 1992
Db 3937 TACGGACGAACAGGAGGATCCGTGTCCTCTCTTTTAAAACTGGCATCATTTCCCTGTG 3996
Qy 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 3997 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4056
Qy 2053 ATTTTGTGACCGAGCGAGCTGGGCTCTCTTCTGTCATGATTCTATCCAA 2101
Db 4057 ATTTTGTGACCGAGCGAGCTGGGCTCTCTTCTGTCATGATTCTATCCAA 4105

RESULT 8
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.
XX
PS Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a
cytomegalovirus (CMV) promoter and a small polyA signal sequence

5Q	Sequence	4848 BP;	1405 A;	1144 C;	1180 G;	1119 T;	0 U;	0 Other;	
	Query Match	83.0%;	Score	1743;	DB	6;	Length	4848;	
	Best local Similarity	85.8%;	Pred. No.	0;					
	Matches 2101;	Conservative	0;	Mismatches	0;	Indels	348;	Gaps	1;
2Y	1	GAGCTATGCTTACACACACAGGCTGTTATGTCAACCACCTCTGTACCCCTACCGGAGCCCATTT	60						
Db	1680	GAGCTATGCTTACACACACAGGCTGTTATGTCAACCACCTCTGTACCCCTACCGGAGCCCATTT	1739						
2Y	61	TCCTTCAACACCATTTGGAAGCTCTCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG	120						
Db	1740	TCCTTCAACACCATTTGGAAGCTCTCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG	1799						
2Y	121	TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTTC	180						
Db	1800	TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTTC	1859						
QY	181	TGCTGAGGACACATTGCAAGCACAGGAGAGATTTTCTTAATGATGTGGAAGTGGTGAAGA	240						
Db	1860	TGCTGAGGACACATTGCAAGCACAGGAGAGATTTTCTTAATGATGTGGAAGTGGTGAAGA	1919						
QY	241	CCAGTTTCTATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAAGGCCGGGTTGG	300						
Db	1920	CCAGTTTCTATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAAGGCCGGGTTGG	1979						
QY	301	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA	360						
Db	1980	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA	2039						
QY	361	AAC TGAAGTACAAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC	420						
Db	2040	AAC TGAAGTACAAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC	2099						
QY	421	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	480						
Db	2100	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	2159						
QY	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA	540						
Db	2160	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGAAGA	2219						
QY	541	GCCTCTTGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAAATAGGTGCTTCA	600						
Db	2220	GCCTCTTGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAAATAGGTGCTTCA	2279						
QY	601	AGAAGATCTAGAACAAAGAACAAAGTCAGGCTCAATCTCTCACTCACATGGTGGGTAGT	660						
Db	2280	AGAAGATCTAGAACAAAGAACAAAGTCAGGCTCAATCTCTCACTCACATGGTGGGTAGT	2339						
QY	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	720						
Db	2340	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	2399						
QY	721	AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAAGA---	777						
Db	2400	AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAAGACAC	2459						
QY	778	-----	777						
Db	2460	TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGA	2519						
QY	778	-----	777						
Db	2520	AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAGA	2579						
QY	778	-----	777						
Db	2580	CTCCAAGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTTCCAAGGTGAAATGAAGC	2639						
QY	778	-----	777						
Db	2640	TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA	2699						

QY	778	-----	777						
Db	2700	AGGTTCCGATGATGCAGTCCTGTGTACAAAGACGTTTGGATAACATGAACCTTCAACTGGAG	2759						
QY	778	-----	792						
Db	2760	TGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG	2819						
QY	793	GAAGCGTCTGCACCTTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA	852						
Db	2820	GAAGCGTCTGCACCTTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA	2879						
QY	853	ATTAAGCCCGCAGGCACCTATTGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT	912						
Db	2880	ATTAAGCCCGCAGGCACCTATTGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT	2939						
QY	913	ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA	972						
Db	2940	ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA	2999						
QY	973	GACTGTACGAATATTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA	1032						
Db	3000	GACTGTACGAATATTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA	3059						
QY	1033	GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAGCA	1092						
Db	3060	GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAGCA	3119						
QY	1093	GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTCGCTCCGCTGACTGGCAGAG	1152						
Db	3120	GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTCGCTCCGCTGACTGGCAGAG	3179						
QY	1153	AAAAATAGATGAGACCCCTTGAAAGACTCCAGAGAACTTCAAGAGGCCACCGATGAGCTGGA	1212						
Db	3180	AAAAATAGATGAGACCCCTTGAAAGACTCCAGAGAACTTCAAGAGGCCACCGATGAGCTGGA	3239						
QY	1213	CCTCAAGCTCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT	1272						
Db	3240	CCTCAAGCTCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT	3299						
QY	1273	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAAAATTGGGCC	1332						
Db	3300	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAAAATTGGGCC	3359						
QY	1333	TCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT	1392						
Db	3360	TCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT	3419						
QY	1393	TCAGCTCTCACCCGTATAACCTCAGCACCTCTGGAAGACCTGAACACACAGATGGAAGCTTCT	1452						
Db	3420	TCAGCTCTCACCCGTATAACCTCAGCACCTCTGGAAGACCTGAACACACAGATGGAAGCTTCT	3479						
QY	1453	GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCC	1512						
Db	3480	GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCC	3539						
QY	1513	AGCATCTCAGCACCTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC	1572						
Db	3540	AGCATCTCAGCACCTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC	3599						
QY	1573	AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGTGGGACCATCCCAA	1632						
Db	3600	AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGTGGGACCATCCCAA	3659						
QY	1633	AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAAGTCTCAGCTTATAG	1692						
Db	3660	AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAAGTCTCAGCTTATAG	3719						
QY	1693	GACTGCCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	1752						
Db	3720	GACTGCCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	3779						

1753 AGCTGATGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCATGGATAT 1812
|||||
3780 AGCTGATGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCATGGATAT 3839
|||||
1813 CCTGCAGATTATTAATTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCAACAACAA 1872
|||||
3840 CCTGCAGATTATTAATTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCAACAACAA 3899
|||||
1873 TTTGGTCAACGTCCTCTCTCGGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGA 1932
|||||
3900 TTTGGTCAACGTCCTCTCTCGGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGA 3959
|||||
1933 TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTG 1992
|||||
3960 TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTG 4019
|||||
1993 TAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
|||||
4020 TAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079
|||||
2053 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCAA 2101
|||||
4080 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCAA 4128
|||||

ESULT 9
AD37264

D AAD37264 standard; DNA; 5060 BP.

C AAD37264;

X

T 21-AUG-2002 (first entry)

E Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.

X Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
X adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
X Becker muscular dystrophy; ds.

S Homo sapiens.
S Cytomegalovirus.
S Unidentified.
S Chimeric.

X WO200183695-A2.

N 08-NOV-2001.

X 27-APR-2001; 2001WO-US013677.

F 28-APR-2000; 2000US-0200777P.

X (XIAO/) XIAO X.

A Xiao X;

I X

X WPI; 2002-049342/06.

R New dystrophin minigene for treating Duchenne or Becker muscular
X dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.

T Example 1; Page 70-71; 71pp; English.

S The present invention relates to an isolated nucleotide sequence encoding
X a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
CC small polyA signal sequence

XX Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

SQ Query Match 83.0%; Score 1743; DB 6; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
|||||
Db 1892 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 1951
|||||
QY 61 TCCTTACACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
|||||
Db 1952 TCCTTACACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 2011
|||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCCTTC 180
|||||
Db 2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCCTTC 2071
|||||
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
Db 2072 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131
|||||
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
|||||
Db 2132 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 2191
|||||
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 360
|||||
Db 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 2251
|||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2311
|||||
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
Db 2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2371
|||||
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAGGAAAAATGGAGGAAGA 540
|||||
Db 2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAGGAAAAATGGAGGAAGA 2431
|||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 600
|||||
Db 2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 2491
|||||
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAGTGGTGGTAGT 660
|||||
Db 2492 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAGTGGTGGTAGT 2551
|||||
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACCACTTAAGGTATTGGG 720
|||||
Db 2552 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACCACTTAAGGTATTGGG 2611
|||||
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGA--- 777
|||||
Db 2612 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACAC 2671
|||||
QY 778 ----- 777
Db 2672 TCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAAGTTTCTTGCTGCTTACAGA 2731
|||||
QY 778 ----- 777
Db 2732 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 2791
|||||
QY 778 ----- 777

b 2792 CTCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2851
y 778 ----- 777
b 2852 TCACACAGATGTTTATCACAAACCTGGATGAAACACAGCCAAAAATCCTGAGATCCCTGGA 2911
y 778 ----- 777
b 2912 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2971
y 778 -----CAGTTCTGACCAGTG 792
b 2972 TGAACCTTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 3031
y 793 GAAGCGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
b 3032 GAAGCGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091
y 853 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAGAACGATGT 912
b 3092 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAGAACGATGT 3151
y 913 ACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGACTACTTTGA 972
b 3152 ACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGACTACTTTGA 3211
y 973 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1032
b 3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 3271
y 1033 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAAATGTCACTGGCTTCTACGAAAGCA 1092
b 3272 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAAATGTCACTGGCTTCTACGAAAGCA 3331
y 1093 GGCTGAGGAGGTCAATACTAGTGAGTGGGAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
b 3332 GGCTGAGGAGGTCAATACTAGTGAGTGGGAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG 3391
y 1153 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGGGCCACGGATGAGCTGGA 1212
b 3392 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 3451
y 1213 CCTCAAGCTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1272
b 3452 CCTCAAGCTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3511
y 1273 CATTGACTCTCTCCAAGATCACTCGAGAACTCAAGGCACTTCGAGGAGAAATTTGGGCC 1332
b 3512 CATTGACTCTCTCCAAGATCACTCGAGAACTCAAGGCACTTCGAGGAGAAATTTGGGCC 3571
y 1333 TCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGCAGGCACTTCGCCAGCTTACCACCTTTGGGCAT 1392
b 3572 TCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGCAGGCACTTCGCCAGCTTACCACCTTTGGGCAT 3631
y 1393 TCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 1452
b 3632 TCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 3691
y 1453 GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTGC 1512
b 3692 GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTGC 3751
y 1513 AGCATCTCAGCACTTTCTTTCCACGTCGTGCCAGGTCCTGGGAGAGAGCCATCTCGCC 1572
b 3752 AGCATCTCAGCACTTTCTTTCCACGTCGTGCCAGGTCCTGGGAGAGAGCCATCTCGCC 3811
y 1573 AAACAAAGTCCCTACTATATCAACCAGAGACTCAAAACACTTGTGGGACCATCCCAA 1632
b 3812 AAACAAAGTCCCTACTATATCAACCAGAGACTCAAAACACTTGTGGGACCATCCCAA 3871
y 1633 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692

Db 3872 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGATTCAGATTCTCAGCTTATAG 3931
QY 1693 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGTCTTGGATCTCTTTGAGCCTGTC 1752
Db 3932 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGTCTTGGATCTCTTTGAGCCTGTC 3991
QY 1753 AGTGCATGTGATGCTTGGACCAGCACACACCTCAAGCAAAATGACAGAGCCCATGGATAT 1812
Db 3992 AGTGCATGTGATGCTTGGACCAGCACACACCTCAAGCAAAATGACAGAGCCCATGGATAT 4051
QY 1813 CCTGCAGATTATTAAATTTGTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAACAA 1872
Db 4052 CCTGCAGATTATTAAATTTGTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAACAA 4111
QY 1873 TTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGTGATGATGTTTATGA 1932
Db 4112 TTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGTGATGATGTTTATGA 4171
QY 1933 TACGGACGAACAGGAGGATCCGTGTCCTCTCTTTTAAAACTGGCATCATTTCCCTGTG 1992
Db 4172 TACGGACGAACAGGAGGATCCGTGTCCTCTCTTTTAAAACTGGCATCATTTCCCTGTG 4231
QY 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4232 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4291
QY 2053 ATTTGTGACGAGCGAGGCTGGGCTCCTCTCTGTCATGATTCTATCCAA 2101
Db 4292 ATTTGTGACGAGCGAGGCTGGGCTCCTCTCTGTCATGATTCTATCCAA 4340

RESULT 10

AAD37260

ID AAD37260 standard; DNA; 4414 BP.

XX

AC AAD37260;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus vector plasmid, AAV-MCK-3447.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

OS Homo sapiens.

OS Unidentified.

OS Chimeric.

XX PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

XX Example 1; Page 65-66; 71pp; English.

PS

XX The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC

domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;

Query Match 78.1%; Score 1641; DB 6; Length 4414;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;

/	1	GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT	60
/	1657	GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT	1716
/	61	TCCTTTCACAGCATTTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG	120
/	1717	TCCTTTCACAGCATTTTGGAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG	1776
/	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	180
/	1777	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC	1836
/	181	TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240
/	1837	TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1896
/	241	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTTACAGCCCATCAGGGCCGGTTGG	300
/	1897	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTTACAGCCCATCAGGGCCGGTTGG	1956
/	301	TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTATCAAGAATGAAGA	360
/	1957	TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTATCAAGAATGAAGA	2016
/	361	AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	420
/	2017	AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	2076
/	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	480
/	2077	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	2136
/	481	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAAGGAAAATGGAGGAAGA	540
/	2137	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAAGGAAAATGGAGGAAGA	2196
/	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAACTACAAACAATAAGGTGCTTCA	600
/	2197	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAACTACAAACAATAAGGTGCTTCA	2256
/	601	AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTAGT	660
/	2257	AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTAGT	2316
/	661	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG	720
/	2317	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG	2376
/	721	AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG	780
/	2377	AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT	2436
/	781	TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT	840
/	2437	CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTCTAGTGCACTGCTTTCAGA	2496
/	841	GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTCAGAA	900

Db	2497	AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAATGTT	2556
Qy	901	GCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAAATCAT	960
Db	2557	ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2616
Qy	961	GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA	1020
Db	2617	GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG	2662
Qy	1021	ACTCTACAGGAGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGACCCAGAAATGTCACTCGGCT	1080
Db	2663	AATAAGTCAGTACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGG--	2720
Qy	1081	TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGGAATAATTGAACCTGCACCTCCGC	1140
Db	2721	-----ATAATTAGTCCAAAAAACTTGAA-----	2743
Qy	1141	TGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCAC	1200
Db	2744	-----AAGAGTACAGCACAGACCCCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCAC	2793
Qy	1201	GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTTGGCAGGCCGT	1260
Db	2794	GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTTGGCAGGCCGT	2853
Qy	1261	GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAAGAAAGTCAAGGCACCTTCGAGG	1320
Db	2854	GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAAGAAAGTCAAGGCACCTTCGAGG	2913
Qy	1321	AGAAATTGCGCCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGTCTGCCAGCTTAC	1380
Db	2914	AGAAATTGCGCCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGTCTGCCAGCTTAC	2973
Qy	1381	CACTTTGGGCAATTCAGCTCTCAACCCTATAACCTCAGCACTCTGGAAGACCTGAACACCAG	1440
Db	2974	CACTTTGGGCAATTCAGCTCTCAACCCTATAACCTCAGCACTCTGGAAGACCTGAACACCAG	3033
Qy	1441	ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCACAG	1500
Db	3034	ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCACAG	3093
Qy	1501	GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGCCAGGGTCCCTGGGAGAG	1560
Db	3094	GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGCCAGGGTCCCTGGGAGAG	3153
Qy	1561	AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG	1620
Db	3154	AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG	3213
Qy	1621	GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGTGCCTGAATAATGTCAGATT	1680
Db	3214	GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGTGCCTGAATAATGTCAGATT	3273
Qy	1681	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGBAGGCCCTTTTGTCTGGATCT	1740
Db	3274	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGBAGGCCCTTTTGTCTGGATCT	3333
Qy	1741	CTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGACTGCAGAACCTCAAGCAAAATGACCA	1800
Db	3334	CTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGACTGCAGAACCTCAAGCAAAATGACCA	3393
Qy	1801	GCCCATGGATATCCTGCAGATTATTAAATTTGTTGACCACCTATTATGACCCCTGGAGCA	1860
Db	3394	GCCCATGGATATCCTGCAGATTATTAAATTTGTTGACCACCTATTATGACCCCTGGAGCA	3453
Qy	1861	AGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGATATGTGTCTGAACCTGGCTGCT	1920
Db	3454	AGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGATATGTGTCTGAACCTGGCTGCT	3513
Qy	1921	GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTTTTAAACCTGGCAT	1980

b 3514 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTTTTAAACTGGCAT 3573
y 1981 CATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
b 3574 CATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 3633
y 2041 AAGTTCACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCA 2100
b 3634 AAGTTCACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCA 3693
y 2101 A 2101
b 3694 A 3694
RESULT 11
AD37242
ID AAD37242 standard; DNA; 3446 BP.
X AAD37242;
X AAD37242;
Y 21-AUG-2002 (first entry)
X Human dystrophin minigene delta3447.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
X adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
X Homo sapiens.
X WO200183695-A2.
PN 08-NOV-2001.
X 27-APR-2001; 2001WO-US013677.
X 28-APR-2000; 2000US-0200777P.
X (XIAO/) XIAO X.
X Xiao X;
X WPI; 2002-049342/06.
X New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
X Example 1; Page 53-54; 71pp; English.
X The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and CR
CC domain) and 11047-11058 (dystrophin last 3 amino acids)
X
X Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 U; 0 Other;
SQ
Query Match 77.5%; Score 1629; DB 6; Length 3446;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 60

Db 900 GAGCTATGCCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCITTCACAGCATTTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTTGAGAG 120
Db 960 TCCITTCACAGCATTTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTTGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATTTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATTTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATCTATGAGGCTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1140 CCAGTTTTCATCTATGAGGCTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAATATTCTCAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTCAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGATCTCCAGAAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGATCTCCAGAAATC- 1378
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
Db 1379 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1438
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 600
Db 1439 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 1498
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTAGT 660
Db 1499 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTAGT 1558
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATGGG 720
Db 1559 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATGGG 1618
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAGACAG 780
Db 1619 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAGACAT 1678
QY 781 TTCTGACCAGTGAAGCGTCTGCACCTTCTCTGAGGAACTTCTGGTGTGGCTACAGCT 840
Db 1679 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 1738
QY 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA 900
Db 1739 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1798
QY 901 GCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAATCAT 960
Db 1799 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1858
QY 961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 1020
Db 1859 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTCTTTTCAACACTGAAG 1904
QY 1021 ACTCTACCAGGAGCCCAAGAGAGCTGCCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCT 1080
Db 1905 AATAAGTCAGTGACCCCAAGAGACGGAAGCATGGCTGGA----- 1942
QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACCTCCGC 1140

b 1943 -----TAACTTTGCCCGTGTGGGATAATTAGTCCAAAAACTTGAA----- 1985
y 1141 TGAATGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACTTCAAGAGGCCAC 1200
b 1986 -----AAGAGTACAGACACAGACCCCTTGAAGACTCCAGAACTTCAAGAGGCCAC 2035
y 1201 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCT 1260
b 2036 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCT 2095
y 1261 GGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
b 2096 GGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2155
y 1321 AGAAATTCGGCCTCTGAAAGAGAACGTCGACACGTCATGACCTTGTCTGCCAGCTTAC 1380
b 2156 AGAAATTCGGCCTCTGAAAGAGAACGTCGACACGTCATGACCTTGTCTGCCAGCTTAC 2215
y 1381 CACTTTGGGCATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
b 2216 CACTTTGGGCATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAG 2275
y 1441 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAG 1500
b 2276 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAG 2335
y 1501 GGACTTTGGTCCAGCATCTCAGCACTTCTTCTCCACGTCCTGTCAGGGTCCCTGGGAGAG 1560
b 2336 GGACTTTGGTCCAGCATCTCAGCACTTCTTCTCCACGTCCTGTCAGGGTCCCTGGGAGAG 2395
y 1561 AGCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAACTTGTG 1620
b 2396 AGCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAACTTGTG 2455
y 1621 GGACATCCCAAAATGACAGAGCTCTACCACTGCTTTAGCTGACCTGAATAATGTCAGAT 1680
b 2456 GGACATCCCAAAATGACAGAGCTCTACCACTGCTTTAGCTGACCTGAATAATGTCAGAT 2515
y 1681 CTCAGCTTATAGGACTGCGCATGAAACTCCGAGACTGCAGAGGCCCTTGTGATGCT 1740
b 2516 CTCAGCTTATAGGACTGCGCATGAAACTCCGAGACTGCAGAGGCCCTTGTGATGCT 2575
y 1741 CTTGAGCCTGTGAGCTGATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACCA 1800
b 2576 CTTGAGCCTGTGAGCTGATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACCA 2635
y 1801 GCCATGGATATCCTGCGAGATTATTAATGTTGACCACTATTTATGACCGCCTGGAGCA 1860
b 2636 GCCATGGATATCCTGCGAGATTATTAATGTTGACCACTATTTATGACCGCCTGGAGCA 2695
y 1861 AGAGCAACAATTTGGTCAACGTCCTCTCTGGGTGGATATGTCGTGAACCTGGCTGCT 1920
b 2696 AGAGCAACAATTTGGTCAACGTCCTCTCTGGGTGGATATGTCGTGAACCTGGCTGCT 2755
y 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCAT 1980
b 2756 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCAT 2815
y 1981 CATTTCCCTGTGTAAGCAACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
b 2816 CATTTCCCTGTGTAAGCAACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2875
y 2041 AAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCCCTCCTTCTGATGATTTCTATCCA 2100
b 2876 AAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCCCTCCTTCTGATGATTTCTATCCA 2935
y 2101 A 2101
b 2936 A 2936

RESULT 12
BK81999

ID XX ABK81999 standard; DNA; 5462 BP.
AC AC ABK81999;
DT DT 13-AUG-2002 (first entry)
XX XX DNA encoding mini-dystrophin protein deltaR2-R21+H3.
DE DE Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX XX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
KW KW Homo sapiens.
OS OS Synthetic.
XX XX WO200229056-A2.
PN PN 11-APR-2002.
XX XX 04-OCT-2001; 2001WO-US031126.
PF PF 06-OCT-2000; 2000US-0238848P.
XX XX (UNMI) UNIV MICHIGAN.
PA PA Chamberlain JS, Harper SQ;
XX XX WPI; 2002-435334/46.
DR DR A composition for preparing therapeutic drugs, has a mini-dystrophin
XX XX peptide comprising a specific number of spectrin-like repeat domains, or
PT PT a nucleic acid sequence encoding the mini-dystrophin peptide.
PT PT Disclosure; Fig 14; 145pp; English.
XX XX The invention describes a composition comprising a mini-dystrophin
CC CC peptide comprising a spectrin-like repeat domain, where the domain
CC CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC CC polynucleotide encoding it is useful as a medicament, for preparing a
CC CC drug for therapeutic application and in the preparation of a composition
CC CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC CC invention
XX XX
SQ SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;
Query Match 77.5%; Score 1629; DB 6; Length 5462;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;
QY 1 GAGCTATGCCTACACACAGSGTGTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGSGTGTATGTACACACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTTGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTTGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278
QY 181 TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360

db 1399 TAATAATTTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1458
2y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 1518
2y 421 TAGCATGAAAAACAAAGCAATTTACATAGATT----- 454
db 1519 TAGCATGAAAAACAAAGCAATTTACATGCTCCTGSACTGACCACCTATTGGAGCCTCTCC 1578
2y 455 ----- 454
db 1579 TACTCAGACTGTTACTCTCTGGTGACACAAACCTGTGGTTTACTAAGGAAACTGCCATCTCCAA 1638
2y 455 -----TTAATGGATCTCCAGAAATCAGAA 477
db 1639 ACTAGAAATGCCATCTTCTCTTGATGTTGGAGCATAGATTACTGCAACAGTTTCCCCCTGGA 1698
2y 478 ACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAATGGAGGA 537
db 1699 CCTGGAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGA 1758
2y 538 AGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACACATAGGTGCT 597
db 1759 TGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACA 1818
2y 598 TCAAGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCAGATGGTGGT 657
db 1819 ATGGCAAGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCAACAACTGGATGA 1878
2y 658 AGTTGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGACA 705
db 1879 AAACAGCCAAAATACTTGAGATCCCTGGAAAGGTTCCGATGATGCAGTCTGTTTACAAG 1938
2y 706 ACTTAAGG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTG 762
db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAG 1998
2y 763 GGTTCCTTTTACAAGACAGTCTTGACCAAGTGGAAAGCTGTGCACCTTTCTCTGCAGGAAC 822
db 1999 GTCCCATTTGGAAAGCCAGTTCTTGACCAAGTGGAAAGCTGTGCACCTTTCTCTGCAGGAAC 2058
2y 823 TCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCGGCGAGGCACCTATTGGAGGCGA 882
db 2059 TCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCGGCGAGGCACCTATTGGAGGCGA 2118
2y 883 CTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAAAC 942
db 2119 CTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAAAC 2178
2y 943 TAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGAGCAGCC 1002
db 2179 TAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGAGCAGCC 2238
2y 1003 TTTGGAAGGACTAGAGAAACTCTACAGGAGGCCAGAGAGCTGCCCTCGAGGAGAGAGC 1062
db 2239 TTTGGAAGGACTAGAGAAACTCTACAGGAGGCCAGAGAGCTGCCCTCGAGGAGAGAGC 2298
2y 1063 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 1122
db 2299 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 2358
2y 1123 ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCA 1182
db 2359 ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCA 2418
2y 1183 GGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAA 1242
db 2419 GGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAA 2478
2y 1243 GGGATCCTTGGCAGCCCCTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 1302
db 2479 GGGATCCTTGGCAGCCCCTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 2538

QY 1303 AGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGA 1362
Db 2539 AGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGA 2598
QY 1363 CTTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCT 1422
Db 2599 CTTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCT 2658
QY 1423 GGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCA 1482
Db 2659 GGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCA 2718
QY 1483 GCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCAGCTCTGT 1542
Db 2719 GCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCAGCTCTGT 2778
QY 1543 CCAGGGTCCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCTACTATATCAACCACGA 1602
Db 2779 CCAGGGTCCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCTACTATATCAACCACGA 2838
QY 1603 GACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGA 1662
Db 2839 GACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGA 2898
QY 1663 CCTGAATAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 1722
Db 2899 CCTGAATAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 2958
QY 1723 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCAGACAA 1782
Db 2959 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCAGACAA 3018
QY 1783 CCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTAT 1842
Db 3019 CCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTAT 3078
QY 1843 TTATGACCGCCTGGAGCAAGAGCAACAATTTGTCACACGTCCCTCTCTGCTGGATAT 1902
Db 3079 TTATGACCGCCTGGAGCAAGAGCAACAATTTGTCACACGTCCCTCTCTGCTGGATAT 3138
QY 1903 GTGTCTGAATGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCT 1962
Db 3139 GTGTCTGAATGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCT 3198
QY 1963 GTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAGATA 2022
Db 3199 GTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAGATA 3258
QY 2023 CTTTTCACAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCT 2082
Db 3259 CTTTTCACAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCT 3318
QY 2083 TCTGCATGATTCTATCCAA 2101
Db 3319 TCTGCATGATTCTATCCAA 3337

RESULT 13

ABK81997

ID ABK81997 standard; DNA; 5417 BP.

XX AC ABK81997;

XX XX

DT 13-AUG-2002 (first entry)

XX XX

DE DNA encoding mini-dystrophin protein deltaR4-R23.

XX XX

XW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

XW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX Homo sapiens.

OS Synthetic.

WO200229056-A2.
11-APR-2002.
04-OCT-2001; 2001WO-US031126.
06-OCT-2000; 2000US-0238848P.
(UNMI) UNIV MICHIGAN.
Chamberlain JS, Harper SQ;
WPI; 2002-435334/46.
A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.
Disclosure; Fig 12; 145pp; English.
The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention
Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;
Query Match 76.7%; Score 1611.6; DB 6; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;
1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 60
1099 GAGCTATGCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 1158
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCCGGTTGG 300
1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCCGGTTGG 1398
301 TAATATTTCAAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
1399 TAATATTTCAAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 420
1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 1518
421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
481 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAACAAAGGAAAATGGAGGAAGA 540
1579 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAACAAAGGAAAATGGAGGAAGA 1638

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCACTAACAACATAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCACTAACAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACAACTCAGGGTCAATTTCTCTCACTACATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACAACTCAGGGTCAATTTCTCTCACTACATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTTGGAAAGAACAACTTAAAGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTTGGAAAGAACAACTTAAAGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAAGAACTTTCTGGTGTGGCTACAGCT 840
Db 1879 CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTGATGGCTTTCAGA 1938
QY 841 GAAAGATGATGAATTAAGCCGGCAGGACACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA 900
Db 1939 AAAAGAGATGTCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 901 GCAGAACGATGTACATAGGGCC--TTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAATC 958
Db 1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCCTTTTGA 1008
Db 2059 GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 2118
QY 1009 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCTCCTGAGGAGAGAGGCCAGAA 1068
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTTAGTCCAAAA 2178
QY 1069 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC----- 1110
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTACAGGCTGTCCACCACTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGACAACCTGTAATGGAACACAGTAACTACGGTGACCACAGGGAACAGATCCTGTGTA 2298
QY 1111 ---TGAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGGAACCTCCACCACCACTCCCCAAAAGAGAGGCGAGATTACTGTGA 2358
QY 1168 CCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCA 1227
Db 2359 TCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCA 2418
QY 1228 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCA 1287
Db 2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCA 2478
QY 1288 AGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGT 1347
Db 2479 AGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGT 2538
QY 1348 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTA 1407
Db 2539 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTA 2598
QY 1408 TAACTCTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGTCTGCAGGTGGCGCTCGA 1467
Db 2599 TAACTCTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGTCTGCAGGTGGCGCTCGA 2658
QY 1468 GGACCGAGTCAAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTT 1527
Db 2659 GGACCGAGTCAAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTT 2718

721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGA --- 777
1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACCA 1679
778 ----- 777
1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCCTCTCTACTCAGACTGTTACTCT 1739
778 ----- 777
1740 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 1799
778 ----- 777
1800 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAA 1859
778 ----- 777
1860 GTTTCCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCG 1919
778 ----- 777
1920 TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 1979
778 ----- 777
1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTTGGATGAAACACAGCCA 2039
778 ----- 777
2040 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGA 2099
778 ----- 777
2100 TAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATT 2159
778 ----- 777
2160 GGAAGCCAGTCTGACCAAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTG 2219
832 GCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGC 891
2220 GCTACAGCTGAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGC 2279
892 AGTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAAC 951
2280 AGTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAAC 2339
952 TGTAATCATGACTACTCTTGAGACTGTACGAATATTTTGACAGAGCAGCCCTTTGGAAG 1011
2340 TGTAATCATGACTACTCTTGAGACTGTACGAATATTTTGACAGAGCAGCCCTTTGGAAG 2399
1012 ACTAGAGAACTCTACGAGAGGCCAGAGAGCTGCCTCTGAGGAGAGAGGCCAGAAATGT 1071
2400 ACTAGAGAACTCTACGAGAGGCCAGAGAGCTGCCTCTGAGGAGAGAGGCCAGAAATGT 2459
1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 1131
2460 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 2519
1132 GCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCA 1191
2520 GCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCA 2579
1192 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTG 1251
2580 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTG 2639
1252 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1311
2640 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 2699
1312 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTTCAATGACCTTGCTCG 1371

Db 2700 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCG 2759
Qy 1372 CCAGCTTACCACCTTTGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCT 1431
Db 2760 CCAGCTTACCACCTTTGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCT 2819
Qy 1432 GAACACCAGATGGAAGCTTCTGCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 1491
Db 2820 GAACACCAGATGGAAGCTTCTGCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2879
Qy 1492 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 1551
Db 2880 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 2939
Qy 1552 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACCAGAGACTCAAAC 1611
Db 2940 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACCAGAGACTCAAAC 2999
Qy 1612 AACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 1671
Db 3000 AACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3059
Qy 1672 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 1731
Db 3060 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 3119
Qy 1732 CTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCCTTGGACGACACAACCTCAAGCA 1791
Db 3120 CTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCCTTGGACGACACAACCTCAAGCA 3179
Qy 1792 AAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATTATGACCG 1851
Db 3180 AAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATTATGACCG 3239
Qy 1852 CCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATATGTGCTGAA 1911
Db 3240 CCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATATGTGCTGAA 3299
Qy 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAA 1971
Db 3300 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAA 3359
Qy 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAGATACCTTTTCAA 2031
Db 3360 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAGATACCTTTTCAA 3419
Qy 2032 GCAAGTGGCAAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCCCTCCTTCTGCAATGA 2091
Db 3420 GCAAGTGGCAAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCCCTCCTTCTGCAATGA 3479
Qy 2092 TTCTATCCAA 2101
Db 3480 TTCTATCCAA 3489

RESULT 15
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS Unidentified.
OS Chimeric.

IX WO200183695-A2.
IX 08-NOV-2001.
IX 27-APR-2001; 2001WO-US013677.
IX 28-APR-2000; 2000US-0200777P.
IX (XIAO/) XIAO X.
IX Xiao X;
IX WPI; 2002-049342/06.
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
IX Example 1; Page 59-60; 71pp; English.
IX The present invention relates to an isolated nucleotide sequence encoding
IX a dystrophin minigene. The minigene comprises N-terminal or modified N-
IX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
IX domains and cysteine-rich domains of dystrophin or utrophin genes. The
IX invention also relates to a recombinant adeno-associated virus (AAV)
IX comprising dystrophin minigene operably linked to an expression control
IX element. The dystrophin minigene in operable linkage with an expression
IX control element, in a recombinant adeno-associated virus or retrovirus is
IX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
IX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
IX vector plasmid construct containing human dystrophin minigenes, a muscle
IX creatine kinase (MCK) promoter and a small polyA signal sequence
IX Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;
IX
IX Query Match 76.2%; Score 1602; DB 6; Length 4966;
IX Best Local Similarity 81.1%; Pred No. 0;
IX Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;
IX
IX 1 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTGACCTACACGGAGCCCAT 60
IX 1657 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTGACCTACACGGAGCCCAT 1716
IX
IX 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG 120
IX 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG 1776
IX
IX 121 TGAAGTAACTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTTC 180
IX 1777 TGAAGTAACTGGACCGTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTTC 1836
IX
IX 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTCGTGAAAGA 240
IX 1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTTAATGATGTGGAAGTCGTGAAAGA 1896
IX
IX 241 CCAGTTTTCATACTCAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 300
IX 1897 CCAGTTTTCATACTCAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 1956
IX
IX 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
IX 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016
IX
IX 361 AACTGAAGTACAAAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
IX 2017 AACTGAAGTACAAAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076
IX
IX 421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 480
IX 2077 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 2136
IX

QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAAGGAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAAGGAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 660
Db 2257 AGAAGATCTAGAAACAAGAACCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGA--- 777
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACCA 2436
QY 778 ----- 777
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCCTCTCTACTCAGACTGTACTCT 2496
QY 778 ----- 777
Db 2497 GGTGACACAACTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 2556
QY 778 ----- 777
Db 2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCCTGGAAAA 2616
QY 778 ----- 777
Db 2617 GTTTCTTGCTGGCTTACAGAAAGTGAACAACTGCCAATGTCTCTACAGGATGCTACCCG 2676
QY 778 ----- 777
Db 2677 TAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 2736
QY 778 ----- 777
Db 2737 CCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACACAGCCA 2796
QY 778 ----- 777
Db 2797 AAAAACTCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTCAAAGACGTTTGGGA 2856
QY 778 ----- 777
Db 2857 TAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTT 2916
QY 778 -----CAGTTCTGACCCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 831
Db 2917 GGAAGCCAGTTCTGACCCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 2976
QY 832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 891
Db 2977 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 3036
QY 892 AGTTCAGAAGCAGAACCGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAAC 951
Db 3037 AGTTCAGAAGCAGAACCGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAAC 3096
QY 952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGG 1011
Db 3097 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGG 3156
QY 1012 ACTAGAGAACTCTACAGGAGCCCGAGAGAGTGCCTCTCTGAGGAGAGAGCCCGAGATGT 1071
Db 3157 ACTAGAGAACTCTACAGGAGCCCGAGAGAGTGCCTCTCTGAGGAGAGAGCCCGAGATGT 3216
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGAAAAAATTGAACCT 1131

Job time : 574.141 secs

||||| 3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCT 3276
||||| 1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACCTTCA 1191
||||| 3277 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACCTTCA 3336
||||| 1192 AGAGGCCACGGATGAGTGGAACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTG 1251
||||| 3337 AGAGGCCACGGATGAGTGGAACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTG 3396
||||| 1252 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGC 1311
||||| 3397 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGC 3456
||||| 1312 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCCA CGTCAATGACCTTGCTCG 1371
||||| 3457 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCCA CGTCAATGACCTTGCTCG 3516
||||| 1372 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 1431
||||| 3517 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 3576
||||| 1432 GAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGA 1491
||||| 3577 GAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGA 3636
||||| 1492 AGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCCTCCAGGGTCC 1551
||||| 3637 AGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCCTCCAGGGTCC 3696
||||| 1552 CTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATATCAACCACGAGACTCAAAC 1611
||||| 3697 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 3756
||||| 1612 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCTGACCTGAATAA 1671
||||| 3757 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCTGACCTGAATAA 3816
||||| 1672 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTG 1731
||||| 3817 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTG 3876
||||| 1732 CTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACAACTCAAGCA 1791
||||| 3877 CTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACAACTCAAGCA 3936
||||| 1792 AAATGACCAAGCCCATGGATATCCTGACATTATTAATTGTTGACCACTATTTATGACCG 1851
||||| 3937 AAATGACCAAGCCCATGGATATCCTGACATTATTAATTGTTGACCACTATTTATGACCG 3996
||||| 1852 CCTGGAGCAAGAGACAAACAATTGTTGTAACGTCCTCTCTGCGTGGATATGTGCTGAA 1911
||||| 3997 CCTGGAGCAAGAGACAAACAATTGTTGTAACGTCCTCTCTGCGTGGATATGTGCTGAA 4056
||||| 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTTTTAA 1971
||||| 4057 CTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTTTTAA 4116
||||| 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 2031
||||| 4117 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 4176
||||| 2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGCAATGA 2091
||||| 4177 GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGCAATGA 4236
||||| 2092 TTCTATCCAA 2101
||||| 4237 TTCTATCCAA 4246

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 nucleic - nucleic search, using sw model
in on: April 4, 2004, 12:04:07 ; Search time 5477.46 Seconds
(without alignments)
16625.161 Million cell updates/sec
tle: US-09-845-416-12_COPY_900_3000
rfect score: 2101
equence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101
oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
eached: 3470272 seqs, 21671516995 residues
otal number of hits satisfying chosen parameters: 6940544
inimum DB seq length: 0
aximum DB seq length: 2000000000
ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1760.4	83.8	5339	6	AX538620	AX538620 Sequence
2	1629	77.5	5462	6	AX538621	AX538621 Sequence
3	1611.6	76.7	5417	6	AX538619	AX538619 Sequence
4	1328.8	63.2	5952	6	AR304538	AR304538 Sequence
5	1328.8	63.2	5952	6	AX114289	AX114289 Sequence
6	1328.8	63.2	8689	6	AX538622	AX538622 Sequence
7	1328.8	63.2	11443	6	AX538624	AX538624 Sequence
8	1328.8	63.2	12057	6	AX538627	AX538627 Sequence
9	1328.8	63.2	12446	9	HSDMDR	X14298 Human mRNA
10	1328.8	63.2	13957	6	AX409637	AX409637 Sequence
11	1328.8	63.2	13957	6	AX538581	AX538581 Sequence
12	1328.8	63.2	13957	9	HUMDYS	M18533 Homo sapien
13	1317.8	62.7	13977	6	AR220819	AR220819 Sequence
14	1317	62.7	4402	6	E30219	E30219 Shortened d
15	1196.6	57.0	4402	6	E30220	E30220 Shortened d
16	1185.4	56.4	13887	4	AF070485	AF070485 Canis fam
17	1149.6	54.7	13815	6	AX306153	AX306153 Sequence
18	1149.6	54.7	13815	6	AX538582	AX538582 Sequence
19	1149.6	54.7	13815	10	MUSDYSA	M68859 Mouse dyster
20	1149.6	54.7	19307	6	AR093392	AR093392 Sequence
21	1149.6	54.7	19307	6	AR142592	AR142592 Sequence
22	922.2	43.9	13575	5	GGDYS	X13369 Chicken mRN
23	718.8	34.2	4075	6	E30221	E30221 Shortened d
24	715.6	34.1	2654	5	FSCDYSTRO	M37645 Torpedo cal
25	667.2	31.8	3747	6	E30218	E30218 Shortened d
26	629.8	30.0	3275	10	MUSDYS	M18025 Mouse dyster
27	613.6	29.2	3163	6	E30223	E30223 Shortened d
28	598.2	28.5	3521	5	AF339031	AF339031 Danio rer
29	555.6	26.4	10705	10	RNAJ2967	AJ002967 Rattus no
30	554.2	26.4	11096	6	AX538584	AX538584 Sequence
31	554.2	26.4	11096	10	MMY12229	Y12229 M.musculus
32	554	26.4	3161	10	MMGUTRPH	X83506 M.musculus
33	547.4	26.1	6045	6	A63605	A63605 Sequence 7
34	547.4	26.1	6045	6	AR281528	AR281528 Sequence
35	547.4	26.1	6059	6	AX107972	AX107972 Sequence
36	547.4	26.1	10302	6	AX538583	AX538583 Sequence
37	547.4	26.1	10302	9	HSMUPS	X69086 H.sapiens m
38	547.4	26.1	10320	6	A63607	A63607 Sequence 9
39	547.4	26.1	10320	6	AR281529	AR281529 Sequence
40	528.4	25.1	3073	10	AF195788	AF195788 Rattus no
41	528.4	25.1	3140	10	AF195787	AF195787 Rattus no
42	527.4	25.1	3499	9	HSU43519	U43519 Human dyster
43	527.4	25.1	5106	6	AX552248	AX552248 Sequence
44	514.6	24.5	1481	9	BC009242	BC009242 Homo sapi
45	505	24.0	2005	6	AX817318	AX817318 Sequence

ALIGNMENTS

RESULT 1
AX538620
LOCUS AX538620
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers
source	1..5339	/organism="synthetic construct"
		/mol_type="unassigned DNA"
		/db_xref="taxon:32630"
		/note="Synthetic"
ORIGIN		
Query Match 83.8%; Score 1760.4; DB 6; Length 5339;		
Best Local Similarity 90.3%; Pred. No. 0;		
Matches 1910; Conservative 0; Mismatches 191; Indels 15; Gaps 2;		
1	1	GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGTACCCCTACACGGAGCCCAT 60
1099		GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 1158
61		TCCTTCACAGCATTGTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCAATGATGGAGAG 120
1159		TCCTTCACAGCATTGTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCAATGATGGAGAG 1218
121		TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
1219		TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278
181		TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240
1279		TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1338
241		CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
1339		CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
301		TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAAGAGAAATTTATCAGAAGATGAAGA 360
1399		TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAAGAGAAATTTATCAGAAGATGAAGA 1458
361		AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
1459		AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
421		TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
1519		TAGCATGGAAAAACAAAGCAATTTACATCATAGATTACTGCAACAGTCCCTGGACCT 1578
481		GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAAAGGAAATGGAGGAAGA 540
1579		GGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGC 1638
541		GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTAGTACAACATAAGGTGCTTCA 600
1639		TACCCGTAAAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATG 1698
601		AGAAGATCTAGAACACAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 660
1699		GCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAA 1758
661		TGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACT 708
1759		CAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACG 1818
709		TAAGG---TATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGTGGGT 765
1819		TTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTC 1878
766		TCCTTTACAAGACAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTCT 825
1879		CCATTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTCT 1938
826		GGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTT 885
1939		GGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTT 1998
886		TCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAA 945

Db	1999	TCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAA 2058
QY	946	AGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 1005
Db	2059	AGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 2118
QY	1006	GGAAGGACTAGAGAAAACCTCTACCAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCA 1065
Db	2119	GGAAGGACTAGAGAAAACCTCTACCAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCA 2178
QY	1066	GAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATT 1125
Db	2179	GAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATT 2238
QY	1126	GAACTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGA 1185
Db	2239	GAACTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGA 2298
QY	1186	ACTTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGG 1245
Db	2299	ACTTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGG 2358
QY	1246	ATCCTGGCAGCCCGTGGGCGATCTCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGT 1305
Db	2359	ATCCTGGCAGCCCGTGGGCGATCTCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGT 2418
QY	1306	CAAGGCACTTCGAGGAGAAATTGCGCCTCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCT 1365
Db	2419	CAAGGCACTTCGAGGAGAAATTGCGCCTCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCT 2478
QY	1366	TGCTCGCCAGCTTACCCTTTGGGCATTTGAGCTCTCACCGTATAACCTCAGCACTCTGGA 1425
Db	2479	TGCTCGCCAGCTTACCCTTTGGGCATTTGAGCTCTCACCGTATAACCTCAGCACTCTGGA 2538
QY	1426	AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAGGCAGCT 1485
Db	2539	AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAGGCAGCT 2598
QY	1486	GCATGAAGCCCAAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCA 1545
Db	2599	GCATGAAGCCCAAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCA 2658
QY	1546	GGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATATCAACCCACGAGAC 1605
Db	2659	GGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATATCAACCCACGAGAC 2718
QY	1606	TCAAACAACCTGTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1665
Db	2719	TCAAACAACCTGTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2778
QY	1666	GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGC 1725
Db	2779	GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGC 2838
QY	1726	CCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCCTTGGACCAGCACAACT 1785
Db	2839	CCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCCTTGGACCAGCACAACT 2898
QY	1786	CAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTA 1845
Db	2899	CAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTA 2958
QY	1846	TGACCCGCTGGAGCAAGAGCAACAATTTGGTCAAGTGCCTCTCTGCGTGGATATGTG 1905
Db	2959	TGACCCGCTGGAGCAAGAGCAACAATTTGGTCAAGTGCCTCTCTGCGTGGATATGTG 3018
QY	1906	TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTCTCTGTC 1965
Db	3019	TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTCTCTGTC 3078
QY	1966	TTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAGATACCT 2025

b 3079 TTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACCT 3138
Y 2026 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCGAGGCTGGGCTCCTTCT 2085
b 3139 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCGAGGCTGGGCTCCTTCT 3198
Y 2086 GCATGATTCTATCCAA 2101
b 3199 GCATGATTCTATCCAA 3214

RESULT 2
K538621
JCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
SEQUENCE 41 from Patent WO0229056.
CCESION AX538621
ERSION AX538621.1 GI:25271168
EYWORDS
OURCE synthetic construct
OURGEM synthetic construct
artificial sequences.

EFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

EATURES Location/Qualifiers
source 1. 5462
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

RIGIN

Query Match 77.5%; Score 1629; DB 6; Length 5462;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;

Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
D 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158
Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
D 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
Y 121 TGAAGTAAACCTGGACCGTTATCAACAGCCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
D 1219 TGAAGTAAACCTGGACCGTTATCAACAGCCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278
Y 181 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 240
D 1279 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 1338
Y 241 CCAGTTTCATATCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
D 1339 CCAGTTTCATATCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1398
Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
D 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
D 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGATT----- 454
D 1519 TAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTGACCACTATTGGAGCCTCTCC 1578
Y 455 ----- 454
D 1579 TACTCAGACTGTTACTCTGGTGACAAACCTGTGGTTACTAAGGAAACTGCCATCTCCAA 1638

QY 455 -----TTAATGGATCTCCAGAAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCTTGTGATGTTGGAGCATAGATTACTGCAACAGTTCCTCCCTGGA 1698
QY 478 ACTGAAAGAGTTGATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGA 537
Db 1699 CCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGA 1758
QY 538 AGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTCAACAAACATAAGGTGCT 597
Db 1759 TGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACA 1818
QY 598 TCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGT 657
Db 1819 ATGGCAAGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGA 1878
QY 658 AGTTGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACA 705
Db 1879 AAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTTTACAAAG 1938
QY 706 ACTTAAGG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTG 762
Db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAG 1998
QY 763 GGTTCCTTTTACAAGACAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACT 822
Db 1999 GTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACT 2058
QY 823 TCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 882
Db 2059 TCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 2118
QY 883 CTTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAAAC 942
Db 2119 CTTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAAAC 2178
QY 943 TAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 1002
Db 2179 TAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 2238
QY 1003 TTTGGAAGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCTCTGAGGAGAGAGC 1062
Db 2239 TTTGGAAGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCTCTGAGGAGAGAGC 2298
QY 1063 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 1122
Db 2299 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 2358
QY 1123 ATTGAACCTGCATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCA 1182
Db 2359 ATTGAACCTGCATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCA 2418
QY 1183 GGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAA 1242
Db 2419 GGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAA 2478
QY 1243 GGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 1302
Db 2479 GGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 2538
QY 1303 AGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAAGAGAACTGAGGCCACCTCAATGA 1362
Db 2539 AGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAAGAGAACTGAGGCCACCTCAATGA 2598
QY 1363 CTTTGTCTGCCAGCTTACCCTTTGGGCAATTGAGCTCTCACCGTATAACTCAGCACTCT 1422
Db 2599 CTTTGTCTGCCAGCTTACCCTTTGGGCAATTGAGCTCTCACCGTATAACTCAGCACTCT 2658
QY 1423 GGAAGACCTGAACACAGATGGAAGCTTTGCGGCAATTGAGCTCTCACCGTATAACTCAGCACTCT 1482
Db 2659 GGAAGACCTGAACACAGATGGAAGCTTTGCGAGGTGGCCGTCGAGGACCGAGTCAGGCA 2718

1483 GCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGT 1542
2719 GCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGT 2778
1543 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGA 1602
2779 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGA 2838
1603 GACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGA 1662
2839 GACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGA 2898
1663 CCTGAATTAATGTGAGATTTCTGAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 1722
2899 CCTGAATTAATGTGAGATTTCTGAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 2958
1723 GGCCCTTTGCTTGGATCTCTGAGCCCTGTGAGTGCATGTGATGCTTGGACCAACAA 1782
2959 GGCCCTTTGCTTGGATCTCTGAGCCCTGTGAGTGCATGTGATGCTTGGACCAACAA 3018
1783 CCTCAAGCAAAATGACAGCCCATGGATATCTGCGAGATTTAATTTGTTGACCACTAT 1842
3019 CCTCAAGCAAAATGACAGCCCATGGATATCTGCGAGATTTAATTTGTTGACCACTAT 3078
1843 TTATGACCCGCTGGAGCAAGAGCAACAAATTTGGTCAAAGTCCCTCTCTGCGTGGATAT 1902
3079 TTATGACCCGCTGGAGCAAGAGCAACAAATTTGGTCAAAGTCCCTCTCTGCGTGGATAT 3138
1903 GTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCT 1962
3139 GTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCT 3198
1963 GTCTTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTTGAAGACAAAGTACAGATA 2022
3199 GTCTTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTTGAAGACAAAGTACAGATA 3258
2023 CCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGAGGCTGGGCTCCT 2082
3259 CCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGAGGCTGGGCTCCT 3318
2083 TCTGATGATTTCTATCCAA 2101
3319 TCTGATGATTTCTATCCAA 3337

RESULT 3
AX538619
LOCUS AX538619 5417 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 39 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. 5417
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN

Query Match 76.7%; Score 1611.6; DB 6; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTGTCACCACCTCTGACCCCTACACGGAGCCCAT 60

Db 1099 GAGCTATGCTTACACACAGGCTGCTTATGTGTCACCACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTTCACAGCAATTTGGAAGCTCTGGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
Db 1159 TCCTTTCACAGCAATTTGGAAGCTCTGGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
QY 121 TGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGCAGGCCCATCAGGCCCGGTTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGCAGGCCCATCAGGCCCGGTTGG 1398
QY 301 TAAATATCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAAATATCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1518
QY 421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 480
Db 1519 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAACAGAAAGAAACAAAGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTTAAACAAACAGAAAGAAACAAAGAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAAATAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAAATAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAGCAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAGCAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAG 780
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAT 1878
QY 781 TTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGTTGGCTACAGCT 840
Db 1879 CTTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 1938
QY 841 GAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAGAA 900
Db 1939 AAAAGAAGATGCAAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 1998
QY 901 GCAGAAGCATGTACATAGGGCC- -TTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAATC 958
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 959 ATGAGTACTCTTGAGACTGTACGAATATTT- - - - -CTGACAGAGCAGCCTTTTGA 1008
Db 2059 GGGCAAACTCTATTCACTCAAAACAGATCTTCTTCAACACTGAAGAATAAGTCAAGTAC 2118
QY 1009 AGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1068
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAA 2178
QY 1069 TGTCACCTCGGCTTCTACGAAGACAGGCTGAGGAGGTCAATAC- - - - - 1110

b 2179 ACTTGAAAGAGTACAGCACAGATTTCACAGGCTGTCAACCACCACTCAGCCATCACTAAC 2238
y 1111 ----- 1110
b 2239 ACAGACAACTGTAATGGAAACAGTAACCTACGGTGACCAACAAGGGAACAGATCCTGGTAAA 2298
y 1111 ---TGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1167
c 2299 GCATGCTCAAGAGGAACCTTCCACCACCTCCCAAAAGAGAGGCAGATTACTGTGGA 2358
y 1168 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 1227
c 2359 TCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 2418
y 1228 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCAATTGACTCTCTCCA 1287
b 2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCAATTGACTCTCTCCA 2478
y 1288 AGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGT 1347
b 2479 AGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGT 2538
y 1348 GAGCCACGTCAATGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 1407
b 2539 GAGCCACGTCAATGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 2598
y 1408 TAACTTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGA 1467
b 2599 TAACTTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGA 2658
y 1468 GGACCGAGTCAGGAGCTGCATGAAGGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1527
b 2659 GGACCGAGTCAGGAGCTGCATGAAGGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 2718
y 1528 TCTTTCCACGTCTGCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTA 1587
b 2719 TCTTTCCACGTCTGCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTA 2778
y 1588 CTATATCAACCAAGAGACTCAAAACAACTTGCTGGGAGAGCCATCCCAAAATGACAGAGCTCTA 1647
b 2779 CTATATCAACCAAGAGACTCAAAACAACTTGCTGGGAGAGCCATCCCAAAATGACAGAGCTCTA 2838
y 1648 CCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACT 1707
b 2839 CCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACT 2898
y 1708 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGC 1767
b 2899 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGC 2958
y 1768 CTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTA 1827
b 2959 CTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTA 3018
y 1828 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCC 1887
b 3019 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCC 3078
y 1888 TCTCTGCGTGATATGTGTCTGAATGCGCTGTGAATGTTTATGATACGGGACGAACAGG 1947
b 3079 TCTCTGCGTGATATGTGTCTGAATGCGCTGTGAATGTTTATGATACGGGACGAACAGG 3138
y 1948 GAGGATCCGTGCTCTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACTTTTGA 2007
b 3139 GAGGATCCGTGCTCTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACTTTTGA 3198
y 2008 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGG 2067
b 3199 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGG 3258
y 2068 CAGGCTGGCCCTCCTTCTGCATGATTCTATCCAA 2101
b 3259 CAGGCTGGCCCTCCTTCTGCATGATTCTATCCAA 3292

RESULT 4

AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
 1..5952
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTTCTTTTACAAGACAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
Db |||||
QY 3284 GGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTCTCTGCAGGAAC 3343
Db |||||
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
Db |||||
QY 3344 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 3403
Db |||||
QY 882 ACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
Db |||||
QY 3404 ACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 3463
Db |||||
QY 942 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db |||||
QY 3464 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3523
Db |||||
QY 1002 CTTTGGAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG 1061
Db |||||
QY 3524 CTTTGGAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG 3583
Db |||||
QY 1062 CCCAGAAATGTACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
Db |||||
QY 3584 CCCAGAAATGTACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 3643
Db |||||
QY 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 1181
Db |||||
QY 3644 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 3703
Db |||||
QY 1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 1241
Db |||||
QY 3704 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 3763
Db |||||
QY 1242 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
Db |||||
QY 3764 AGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 3823
Db |||||
QY 1302 AAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATG 1361
Db |||||
QY 3824 AAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATG 3883
Db |||||
QY 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCAACGTATAA CCTCAGCACTC 1421
Db |||||
QY 3884 ACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCAACGTATAA CCTCAGCACTC 3943
Db |||||
QY 1422 TGAAGACCTGAACACCAAGATGGAAGCTTTGCAGGTGGCCCTGCGAGGACCGAGTCAGGC 1481
Db |||||

```
.b 3944 TGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCAGAGACCGAGTCAGGC 4003
.y 1482 AGTCGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTCGTG 1541
.b 4004 AGTCGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTCGTG 4063
.y 1542 TCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGG 1601
.b 4064 TCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGG 4123
.y 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661
.b 4124 AGACTCAAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAAGTCTTTAGCTG 4183
.y 1662 ACCTGAATAATGTCAGATCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 1721
.b 4184 ACCTGAATAATGTCAGATCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 4243
.y 1722 AGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCGACACA 1781
.b 4244 AGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCGACACA 4303
.y 1782 ACCTCAAGCAAAATGACAGAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCCACTA 1841
.b 4304 ACCTCAAGCAAAATGACAGAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCCACTA 4363
.y 1842 TTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCTCTCGGTGGATA 1901
.b 4364 TTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCTCTCGGTGGATA 4423
.y 1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 1961
.b 4424 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 4483
.y 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
.b 4484 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 4543
.y 2022 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 2081
.b 4544 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 4603
.y 2082 TTTCTGCATGATTCTATCCAA 2101
.b 4604 TTTCTGCATGATTCTATCCAA 4623
```

```
RESULT 5
X114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced
peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
source
1. .5952
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 2897. .2898
/feature
misc_feature 3198. .3199
/feature
```

```
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
/feature="S2 junction site"
QY 762 GGGTCTCTTTTACAAGACAGCTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
Db |||
3284 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 3343
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 881
Db |||
3344 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 3403
QY 882 ACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAA 941
Db |||
3404 ACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAA 3463
QY 942 CTAAAGAACTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db |||
3464 CTAAAGAACTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3523
QY 1002 CTTTGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCTGAGGAGAGAG 1061
Db |||
3524 CTTTGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCTGAGGAGAGAG 3583
QY 1062 CCCAGAACTGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 1121
Db |||
3584 CCCAGAACTGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 3643
QY 1122 AATTGAACCTGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 1181
Db |||
3644 AATTGAACCTGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 3703
QY 1182 AGGAACCTCAAGAGGCCACGATGAGCTGGACCTCAAGCTCGCCAGCTGAGGTGATCA 1241
Db |||
3704 AGGAACCTCAAGAGGCCACGATGAGCTGGACCTCAAGCTCGCCAGCTGAGGTGATCA 3763
QY 1242 AGGGATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
Db |||
3764 AGGGATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 3823
QY 1302 AAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTCAGCCACGTCATG 1361
Db |||
3824 AAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTCAGCCACGTCATG 3883
QY 1362 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db |||
3884 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 3943
QY 1422 TGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
Db |||
3944 TGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 4003
QY 1482 AGCTGCATGAAGCCCAAGGAGGACTTTGGTCCAGCATCTCAGCACTTTTCCACGTCGTG 1541
Db |||
4004 AGCTGCATGAAGCCCAAGGAGGACTTTGGTCCAGCATCTCAGCACTTTTCCACGTCGTG 4063
QY 1542 TCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGC 1601
Db |||
4064 TCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGC 4123
QY 1602 AGACTCAAAACAACTTGTGTTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661
Db |||
4124 AGACTCAAAACAACTTGTGTTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 4183
QY 1662 ACCTGAATAATGTCAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 1721
Db |||
4184 ACCTGAATAATGTCAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 4243
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCGACACA 1781
Db |||
```

b 4244 AGGCCCTTTGCTGGATCTCTTGAGCCTGTCTAGCCTGTCTAGCTGATGTGATGCCCTTGGACCAGCACA 4303
y 1782 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGACAGATTATTAATTTGTGACCACATA 1841
o 4304 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGACAGATTATTAATTTGTGACCACATA 4363
y 1842 TTTATGACCGCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA 1901
o 4364 TTTATGACCGCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA 4423
y 1902 TGTGCTGAACTGGCTGCTGTAATTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCC 1961
o 4424 TGTGCTGAACTGGCTGCTGTAATTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCC 4483
y 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAACTACAGAT 2021
b 4484 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAACTACAGAT 4543
y 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 2081
b 4544 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 4603
y 2082 TTCTGCATGATTTCTATCCAA 2101
b 4604 TTCTGCATGATTTCTATCCAA 4623

ESULT 6
X538622 AX538622 8689 bp DNA linear PAT 23-NOV-2002
OCUS Sequence 42 from Patent WO0229056.
CESSION AX538622
ERSION AX538622.1 GI:25271171
EYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1
EFERENCE Chamberlain,J.S. and Harper,S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 42 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES Location/Qualifiers
source 1. 8689
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

RIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 8689;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
y 762 GGGTTCTTTTACAGACAGTTCTGACCAGTGGAAAGCGTCTGACACCTTTCTCTGAGGAAC 821
b 3330 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGACACCTTTCTCTGAGGAAC 3389
y 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
b 3390 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 3449
y 882 ACTTTCAGCAGTTTCAAGAGCAGAACCATGTACATAGGGCCCTCAAGAGGGGAATTGAAAA 941
b 3450 ACTTTCAGCAGTTTCAAGAGCAGAACCATGTACATAGGGCCCTCAAGAGGGGAATTGAAAA 3509
y 942 CTAAAGAACCTGTAATCATGAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCAGC 1001
b 3510 CTAAAGAACCTGTAATCATGAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCAGC 3569
y 1002 CTTTGGAGGACTACAGAAACTCTACAGAGCCCGCAGAGAGTGCCTCTCTGAGGAGAGAG 1061
b 3570 CTTTGGAGGACTACAGAAACTCTACAGAGCCCGCAGAGAGTGCCTCTCTGAGGAGAGAG 3629

Qy 1062 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 1121
Db 3630 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 3689
Qy 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 1181
Db 3690 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 3749
Qy 1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCA 1241
Db 3750 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCA 3809
Qy 1242 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
Db 3810 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 3869
Qy 1302 AAGTCAAGGCACITTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATG 1361
Db 3870 AAGTCAAGGCACITTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATG 3929
Qy 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCACTC 1421
Db 3930 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCACTC 3989
Qy 1422 TGGAAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGC 1481
Db 3990 TGGAAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGC 4049
Qy 1482 AGCTGCATGAAGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 1541
Db 4050 AGCTGCATGAAGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 4109
Qy 1542 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACG 1601
Db 4110 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACG 4169
Qy 1602 AGACTCAAAACAACCTTGCTGGGACCATCCCAAAAATGACAGAGCTCTACAGTCTTTAGCTG 1661
Db 4170 AGACTCAAAACAACCTTGCTGGGACCATCCCAAAAATGACAGAGCTCTACAGTCTTTAGCTG 4229
Qy 1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
Db 4230 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4289
Qy 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTCTCAGCTGCACTGTGATGTCCTTGACCAGCACA 1781
Db 4290 AGGCCCTTTGCTTGGATCTCTTGAGCCTCTCAGCTGCACTGTGATGTCCTTGACCAGCACA 4349
Qy 1782 ACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCACATA 1841
Db 4350 ACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCACATA 4409
Qy 1842 TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATA 1901
Db 4410 TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATA 4469
Qy 1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1961
Db 4470 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 4529
Qy 1962 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
Db 4530 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 4589
Qy 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 2081
Db 4590 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 4649
Qy 2082 TTCTGCATGATTTCTATCCAA 2101
Db 4650 TTCTGCATGATTTCTATCCAA 4669

Matches 1333;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
762	GGGTTCTTTTACAAGACAGATTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC	821		
8598	GGTCCCAATTGGAAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC	8657		
822	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCG	881		
8658	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCG	8717		
882	ACTTTCAGCAGCTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA	941		
8718	ACTTTCAGCAGCTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA	8777		
942	CTAAAGAACTGTAAATCATAGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGC	1001		
8778	CTAAAGAACTGTAAATCATAGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGC	8837		
1002	CTTTGGAAGGACTAGAGAACTCTACAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAG	1061		
8838	CTTTGGAAGGACTAGAGAACTCTACAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAG	8897		
1062	CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA	1121		
8898	CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA	8957		
1122	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC	1181		
8958	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC	9017		
1182	AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA	1241		
9018	AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA	9077		
1242	AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	1301		
9078	AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	9137		
1302	AAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATG	1361		
9138	AAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATG	9197		
1362	ACCTTGCTCGCAGCTTACCACITTTGGGCATTTACGCTCTCACCGTATAACCTCAGCACTC	1421		
9198	ACCTTGCTCGCAGCTTACCACITTTGGGCATTTACGCTCTCACCGTATAACCTCAGCACTC	9257		
1422	TGGAAGACCTGNAACACCAGATGGAAGCTTCTGCAGGTGSCGTCGAGGACCCAGTCAGGC	1481		
9258	TGGAAGACCTGNAACACCAGATGGAAGCTTCTGCAGGTGSCGTCGAGGACCCAGTCAGGC	9317		
1482	AGCTGCATGAAGCCCAACAGGGACITTTGGTCCAGCATCTCAGCACTTTTCTTTCCACGTCGT	1541		
9318	AGCTGCATGAAGCCCAACAGGGACITTTGGTCCAGCATCTCAGCACTTTTCTTTCCACGTCGT	9377		
1542	TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACG	1601		
9378	TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACG	9437		
1602	AGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG	1661		
9438	AGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG	9497		
1662	ACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCAGA	1721		
9498	ACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCAGA	9557		
1722	AGGCCCTTTGCTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCTTGGACCAGCAC	1781		
9558	AGGCCCTTTGCTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCTTGGACCAGCAC	9617		
1782	ACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA	1841		
9618	ACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA	9677		

Qy	1842	TTTATGACCCCTGGAGCAAGACGACAACAATTGGTCAACGTCCTCTCTCGCTGGATA	1901
Db	9678	TTTATGACCCCTGGAGCAAGACGACAACAATTGGTCAACGTCCTCTCTCGCTGGATA	9737
Qy	1902	TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTGCC	1961
Db	9738	TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTGCC	9797
Qy	1962	TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	2021
Db	9798	TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	9857
Qy	2022	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTTCC	2081
Db	9858	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTTCC	9917
Qy	2082	TTCTGCATGATTCTATCCAA	2101
Db	9918	TTCTGCATGATTCTATCCAA	9937

RESULT 9					
HSDMDR					
LOCUS	HSDMDR	12446 bp	mRNA	linear	PRI 12-SEP-1993
DEFINITION	Human mRNA for dystrophin.				
ACCESSION	X14298				
VERSION	X14298.1	GI:30845			
KEYWORDS	Dmd gene; Duchenne muscular dystrophy; dystrophin.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 12446)				
AUTHORS	Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.				
TITLE	Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous				
JOURNAL	Nucleic Acids Res. 17 (13), 5391 (1989)				
MEDLINE	89345106				
PUBMED	2668885				
REFERENCE	2 (bases 1 to 12446)				
AUTHORS	Rosenthal,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR				
COMMENT	see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.				
FEATURES	Location/Qualifiers				
source	1. .12446				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="X chromosomal, Xp21."				
	/tissue type="muscle"				
	/dev_stage="fetal and adult."				
	99. .11156				
CDS	/note="unnamed protein product; dystrophin (AA 1 - 3685)"				
	/codon_start=1				
	/protein_id="CAA32479.1"				
	/db_xref="GI:30846"				
	/db_xref="GOA:P11532"				
	/db_xref="SWISS-PROT:P11532"				
	/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLPSDL QDGRRLDLLEGLTGQKLPEKSGSTRVHALNNVNKALRVLQNNVDLVNIGSTDIVDG NKLTLGLIWNIIILHWQVNMKNIMAGLOPTNSEKILLSWVRQSTRNYPQVNVINFT TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATORLEHAFNIARYQLGIEKLDPEDV DTTYDPKKSILMYITSLFQVLPQOVSEIAIQEVEMLPRFPVKVKEEHFQLHHQMHYSQ QITVSLAQYVERTSSPKPRFKSVAYTQAAVYVTTSDPTRSPFPFSQHLEAPEDKSFSSS MESEVNLDRYQTALEEVLWLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMLTAM QGRVGNILQLGSKLIGTKLSEDETEVQEQMNLNLSRWECLRVASMEKQSNLHRVLM				

DLNQKLELNDWLTKEERTRKMEBEEPLGPDLEDLKRQVQKHVLOEDLEQEQVRVN	
SLTHMVSWSGSDHATALEQLKVLGDRWANICRWTEADRWVLLQDILLKQWRLTE	
EQCLFSAWLBSEKEDAVNKIHTTGFKDQONEMLSLQKLAVLKADLEKKQSMGLYSIK	
QDLSTLKNKSVTKTQKTEAWLDFNFCWDLNVQLEKSTAQISQAVTTTQPSLTQTITVM	
ETVTTVTTRQILVXHAQEEELPPPPQKKRQITVDSEIRKRLDVIDTELHSHWITRSEA	
VLQSPFAJFRKEGNFSDLKEKVNAIEREKAEKFKLQDASRSQALVEQMVNVEGNA	
DSIQASEQLNSRWIEFQCLLSERLWLEYQNNIIAFYNQLQLEQMTTAAENWLKIQ	
PTTPSEPTAKSQKIKCKEVRNLGSLQPOIERIKIQAISIALKEKGGQPMFLDADFVAF	
TNHFQVFSVDQAREKELQTIPTLPPMRQYQETMSAIRTWQVQSEIKSLIPQSVTDY	
EIMEQRLQALQSSLOEQQGLYILSTTTKEMSKAPSEISRKTSQSEFEIEIGRWK	
KLSSOLVEHCKLEEQMKNLKIQNHIIQTLKKNWAEVDVFLKEEWPALGDSILKKQL	
KQRLLVSDIQTIPSLNSVNEGQKIKNEAPEFASRLKELNLTQWDMCQVY	
ARKEALKGLEKTVSLOKLSMHHEWMTQAEVEYLERDFEYKTPDQLQKAFEEMKRAK	
EEAQKKEAKVLLTESVNSVIAQAPPVAQEAALKKLETLITTYQNLCTRLNGKCKTLE	
EVWACWHELLSYLEKANKWLNVEFKLKTENIPGAEIIESEVLSLENLMRHSBNP	
NQIRILAQITLTDGGVMDLINELETFSNRWRREHAEAVRRQKLEQSIQSAQETENS	
LHLIOESLTFIDKQLAAVIAQVDAQMPQEAQIKQSDLTSHESILEEMKKNQKKEA	
AQVLSQIDVAQKQLQDVSMKFRFQKPFANFEQLQESKMILDEVKXHLPALETKSVE	
QEVVQSLNHCNVNLYKSLSEVKSEVEMVIKTRQIVQKQNTENPKELDERVLTALKHY	
NELGAKVTERKQOLEKCLKLSRKMRKEMNVLTEWLAQDMELTKRSAVEGMPNLDSE	
VAWGKATQKEIEKQKVHLKSTIEVGEALKTVLGKKEITVEDKLSLNSNWIATVSRAE	
EWNLLLEVQKGMETFDQNVHDITKWIIOADTLLDESEKKKQKQEDVLKRLKAELND	
IRPKVDSFTRDQAANLMAHGHCKRLVEPOISELNRFAAISHRKTKGASIPLEKELB	
QFNSDIQKLEPLEAEIQGVNKLKEEDFNKMDNEDNEGTVEKELQKQDNLQQRITDER	
KSEIEKIQQLQTKHNAKDLRSQRRKKALETISHQWYQYKQADLLKCLDDIEKKL	
ALPEPRDERKIKEDRELQKKEELNAVRRQAEGLSEKMIIDVQKQVTEAEQFLKTKQI	
IESKFAQFRLNFAQIHTVREETMMVMTEDMPLEISYVPSYLTITHTVQSALLEVEQ	
LLNAPDLCAKDFEDLFQKEESLKNIKDSLOQSSGRIDIIHKKTAALQASATPVERKL	
QEALSOLDPQWVKVMYKDRQGRFDRSVEKWRPFHYDIKIFQWLTAEAEQFLKTKQI	
PENWEHAKVYKWLKELQDGQKQRTVVRTLNATGEEIIQQSSKTDAASILQEKGLNL	
RWQEVCKQLSDRKRLKEQKILSEFORDLNEFVLEADNIIASIPLEPQKEQQLKE	
KLEQVLLVBEPLRQGLKQNETGGPVLVSAPISPEEQDKLENKQLQTNLQWIKVS	
RALPEKQGEIEAQIKDLGLEKLELLEOLNHLWLSPIRNQLLEIYNQPNQEGPFD	
VKETEIAVQAKQPDVEEILSKGQHLKKEKPAQPVVKRKLDELSSEWKAVNRLQLERA	
KQPDLAGLTTIGASPTQTVTLVTPVVTKETASIKLEMPSSLMLEVPALADFNRAWT	
ELTDWLSLLDQVKSQRMVMDLEDINEMIIKQKATMDQLEQRPQLEELITAAQNLK	
NKTSNOEARTIITDRIERIQNWQDEVEHLQNRQQLNEMLKDSTQWLEAKEEAQVL	
GQARAKLESWKEGPTYVDAIQKKITETKQLAKDLQWQNTVDVANDLAKLLRLDYAD	
DTRKVMITENINASWRSCHKVSEREAALEATHRLLLQFPDLKFLAWLTERETTA	
NVLQDATRKERLLEDKSGVKELMKQWDLQGEIEAHTDVYHNLNDSQKILRSLSGD	
DAVLLORRLDNMFVKSSELRKSLNIRSHLEASDQWKLHLSQELLVWLQKDEL	
SRAPIGGDFPAVQKQNDVHRAPFKRELKTEPVIMSTLETVRJFTEQPLEGLEKLYQ	
EPRELPPEERAQNVTRLRQAEVNTWEKLNLSADQWRKIDLETLEQLQELQBATD	
ELDLKLRQAEVIKGSWQPVGDLIDLSLDHLEKVKALRGEIAPKENVSHVNDLARQL	
TTLIQSLSPYNLSTLEDLNRWKLQVAVEDEVRLQHEAHRDEGPASQHFLESTSVQGP	
WERAISPNKVPYIYNHETQTCWDHPKMTLEYQSLADLNNVRFSAYRTAMKRRILQKA	
LCLDLLSLSAACDALDQHLNKQNDQPMDIQIINCLTTIYDRLEQHNLLVNVPLCVD	
MCLNLLNVYDTRGTRIRVLFSFKTGIISLCKAHLEDKRYLFKQVASSTGFCQRRRL	
GLLLHDSIQIPROLGEVASFGSGNIEPSVSCFPANNKPEIEBAALFLDMRLEPQSM	
WLPVYHRVAAETAKHQAKNICKECPIIGERYRSLKHFNVYDIQSCFFPSRVAKGH	
KMHPVMBYCTPTTSGEDVRDPFAKVLKNKFRKRYFAKHPRMGYLFPVQTVLEGNMET	
PVTLNFPVDSAPASSPOLSHDDTHSRIEHYASRLAEMENSGSYLNDISIPNBSID	
DEHLLIQHYCQSLNQDSPLSQPSPAQILISIESEERLELADLEERENRNLOQAEY	
DRLKQHEHKGSLPLSPPEMPMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDH	
NKQLESQHLRLQLLEQPOAEAKVNGTTVSSPSTSLQSDSSOPMLLRVVGSTSDSM	
GEEDLLSPQDSTSTGLEVNEQLNNSFPSSRGRNTPGKPMREDTM"	

ORIGIN

Query Match	63.2%;	Score	1328.8;	DB	9;	Length	12446;
Best Local Similarity	99.5%;	Pred. No.	0;				
Matches	1333;	Conservative	0;	Mismatches	7;	Indels	0;
						Gaps	0;
QY	762	GGGTTCTTTTACAGACAGTCTGACCAAGTGAAGCGTGTGACCTTTCTCTGCGAGAAC	821				
Db	8488	GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTGTGACCTTTCTCTGCGAGAAC	8547				
QY	822	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGGAGGACCTATTGGAGGCG	881				
Db	8548	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGGAGGACCTATTGGAGGCG	8607				
QY	882	ACTTTCCAGCAGTTCAGAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA	941				
Db	8608	ACTTTCCAGCAGTTCAGAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA	8667				

2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCC 2081
|||||
9748 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCC 9807
|||||
2082 TTCTGCATGATTCTATCCAA 2101
|||||
9808 TTCTGCATGATTCTATCCAA 9827
|||||
SULT 10
409637
CUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
FINITION Sequence 2284 from Patent WO0229103.
CESSION AX409637
RSION AX409637.1 GI:21442342
WORDS
URCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Alves, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
NATURES Location/Qualifiers
source 1. 13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
IGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 13957;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
/ 762 GGGTCTCTTTACAAGACAGTCTGACCCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAAC 821
|||
8598 GGTCCCAATTGGAAGCCAGTTCTGACCCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAAC 8657
|||
822 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCGG 881
|||
8658 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCGG 8717
|||
882 ACTTTCAGCAGTTTCAAGACAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 941
|||
8718 ACTTTCAGCAGTTTCAAGACAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 8777
|||
942 CTAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
|||
8778 CTAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837
|||
1002 CTTTGGAGGACTAGAGAACTCTACCAAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG 1061
|||
8838 CTTTGGAGGACTAGAGAACTCTACCAAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG 8897
|||
1062 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGAAA 1121
|||
8898 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGAAA 8957
|||
1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 1181
|||
8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 9017
|||
1182 AGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGATCA 1241
|||
9018 AGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGATCA 9077
|||
1242 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
|||
9078 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 9137
|||

QY 1302 AAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTCTGAAAGAGAACGCTGAGCCACGTCATG 1361
|||||
Db 9138 AAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTCTGAAAGAGAACGCTGAGCCACGTCATG 9197
|||||
QY 1362 ACCTTGTCTGCCAGCTTACCACTTTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTC 1421
|||||
Db 9198 ACCTTGTCTGCCAGCTTACCACTTTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTC 9257
|||||
QY 1422 TGGAAGACCTGAACACACCCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
|||||
Db 9258 TGGAAGACCTGAACACACCCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGC 9317
|||||
QY 1482 AGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTTCCAGCATCTTCTTCCACGTCTG 1541
|||||
Db 9318 AGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTTCCAGCATCTTCCACGTCTG 9377
|||||
QY 1542 TCCAGGTCTCCCTGGGAGAGAGCATCTCGCCAAACAAAAGTGCCTACTATATCAACCACG 1601
|||||
Db 9378 TCCAGGTCTCCCTGGGAGAGAGCATCTCGCCAAACAAAAGTGCCTACTATATCAACCACG 9437
|||||
QY 1602 AGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
|||||
Db 9438 AGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 9497
|||||
QY 1662 ACCTGAATAATGTCTGAGTTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAGA 1721
|||||
Db 9498 ACCTGAATAATGTCTGAGTTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAGA 9557
|||||
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCACA 1781
|||||
Db 9558 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCACA 9617
|||||
QY 1782 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACTA 1841
|||||
Db 9618 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACTA 9677
|||||
QY 1842 TTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATA 1901
|||||
Db 9678 TTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATA 9737
|||||
QY 1902 TGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTGCC 1961
|||||
Db 9738 TGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTGCC 9797
|||||
QY 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAGAT 2021
|||||
Db 9798 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAGAT 9857
|||||
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCCCTCC 2081
|||||
Db 9858 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCCCTCC 9917
|||||
QY 2082 TTCTGCATGATTCTATCCAA 2101
|||||
Db 9918 TTCTGCATGATTCTATCCAA 9937
|||||

RESULT 11
AX538581
LOCUS Sequence 1 from Patent WO0229056.
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION* AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

linear PAT 23-NOV-2002

FEATURES		Location/Qualifiers	
source	1..13957	/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
RIGIN			
Query Match		63.2%;	Score 1328.8; DB 6; Length 13957;
Best Local Similarity		99.5%;	Pred. No. 0;
Matches 1333;		Conservative 0;	Mismatches 7; Indels 0; Gaps 0;
y	762	GGGTTCTTTTACAAGACAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC	821
db	8598	GGTCCCATTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC	8657
y	822	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG	881
db	8658	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG	8717
y	882	ACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAAA	941
db	8718	ACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAAA	8777
y	942	CTAAGAACCCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC	1001
db	8778	CTAAGAACCCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC	8837
y	1002	CTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG	1061
db	8838	CTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG	8897
y	1062	CCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAA	1121
db	8898	CCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAA	8957
y	1122	AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	1181
db	8958	AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	9017
y	1182	AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCACAGCTGAGGTGATCA	1241
db	9018	AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCACAGCTGAGGTGATCA	9077
y	1242	AGGATCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	1301
db	9078	AGGATCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	9137
y	1302	AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAATG	1361
db	9138	AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAATG	9197
y	1362	ACTTGTCTGCGCAGCTTACCATTGTTGGCATTACCGTCTCACCGTATAACCTCAGCACTC	1421
db	9198	ACTTGTCTGCGCAGCTTACCATTGTTGGCATTACCGTCTCACCGTATAACCTCAGCACTC	9257
y	1422	TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC	1481
db	9258	TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC	9317
y	1482	AGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG	1541
db	9318	AGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG	9377
y	1542	TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACG	1601
db	9378	TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACG	9437
y	1602	AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG	1661
db	9438	AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG	9497
y	1662	ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAGTCCGAGAGACTGCAGA	1721

Db	9498	ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA	9557
QY	1722	AGGCCCTTTTGGATCTCTTGAGCCTGTGACAGCTGCATGTGATGCCTTGGACCAGCACA	1781
Db	9558	AGGCCCTTTTGGATCTCTTGAGCCTGTGACAGCTGCATGTGATGCCTTGGACCAGCACA	9617
QY	1782	ACCTCAAGCAAAATGACCAAGCCCATGGATATGCTGAGATTATTATTTGTTGACCACTA	1841
Db	9618	ACCTCAAGCAAAATGACCAAGCCCATGGATATGCTGAGATTATTATTTGTTGACCACTA	9677
QY	1842	TTTATGACCGCCTGGAGCAAGAGACACAATTTGGTCAACGTCCTCTCTGCGTGGATA	1901
Db	9678	TTTATGACCGCCTGGAGCAAGAGACACAATTTGGTCAACGTCCTCTCTGCGTGGATA	9737
QY	1902	TGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC	1961
Db	9738	TGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC	9797
QY	1962	TGTCCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	2021
Db	9798	TGTCCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	9857
QY	2022	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGTGGGCTCC	2081
Db	9858	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGTGGGCTCC	9917
QY	2082	TTCTGCATGATTCTATCCAA 2101	
Db	9918	TTCTGCATGATTCTATCCAA 9937	
RESULT 12			
HUMDYS	HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000		
LOCUS	Homo sapiens dystrophin (DMD) mRNA, complete cds.		
DEFINITION	M18533 M17154 M18026 M20250		
ACCESSION	M18533.1 GI:181856		
VERSION	Homo sapiens (human)		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1699)		
AUTHORS	Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M.		
TITLE	Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals		
JOURNAL	Cell 50 (3), 509-517 (1987)		
MEDLINE	87273512		
PUBMED	3607877		
REFERENCE	2 (bases 1678 to 3830)		
AUTHORS	Hoffman, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.		
TITLE	Conservation of the Duchenne muscular dystrophy gene in mice and humans		
JOURNAL	Science 238 (4825), 347-350 (1987)		
MEDLINE	88018015		
PUBMED	3659917		
REFERENCE	3 (bases 1 to 13957)		
AUTHORS	Koenig, M., Monaco, A.P. and Kunkel, L.M.		
TITLE	The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein		
JOURNAL	Cell 53 (2), 219-226 (1988)		
MEDLINE	88194521		
PUBMED	3282674		
COMMENT	On May 25, 2000 this sequence version replaced gi:340693. Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.		
FEATURES	Location/Qualifiers		
source	1..13957		

```
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. 13957
/gene="DMD"
209. 11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEEDCYEREDVQKKTFTKWNVAQPSKFGKQHLENLFSDL
QGRRLDLLGLTQGLPKKEKSTGRVHALNNVKNALRVLQNNVDLVNIGSDIVDG
NHKLTGLIWNILHMQVQVMKINIMAGLQQTNSKILLISVQRSTRNPQVNVINFT
TSWSDGLALNALIHSHRPDLFDWNSVVCQSSATQRLHAFNIARYQLGIEKLLDPEDV
DTTPDKKSILMYITSLFQVLPQOVSIEAIQEVEMLPKPKVTKEHFQJHQMHSQ
QITVSLAQGYERTSPKRFKSYATQAYVITSDPTSPSPSQHLEAPEDSKSPGSSL
MESEVNLDRYQTALEVLWSLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLTAH
QQRVGNILQGLSLIGTGLSEDETEYQEMNLNRSWECLRVASMEKQSNLHRLVM
DLQNKQLKELNDLWTKTEERTKMEHEPLGPDLEDLKRQVQKHVLOEDLEGEQVRVN
SLTHWVWVDESSEGDHATAALEEQKVLGDRWANI CRWTEDRWVLLQDILLKWLRLTE
EQCLFSALSEKEDAVNKIHTTGFKDQNMELSSQLKLAVLKADLEKKQSMGLYSLK
QDLSTLKNKSVTKTEAWLDFARNFARNVQKLEKSTAQISQAVTTTPSITQTIVM
ETVTVTTTREQILLVHAQEBELPPPPPKKQITVDSEIRKELDVDTITELHSHWTRSEA
VLQSPFAIFRKEGNFSDLKEKNVNAIEKAEKFRKLQASRSQAQALVQVMNEGVNA
DSIKQASEQLNSRWTFECQLLSERLNLWLEYQNNIIAFYNQLQOLEQMTTAAENWLKIQ
PTTPSEPTAISKQLKICKDEVNRLSGLPQIERKIQISTALKEKQGQPMFLDADFVAF
TNEHFQVFSVDQAREKELQITFDLTPMRYQETMSAIRTWQOQSETKLSIPQLSVTDY
EINEQRLGELQALQSLQEQSGLYLSTTVKEMSKAPSEISIRKYQSEPEIEGRWK
KLSQCLVEHCQKLEQMNKLRKIQNHITQLKMAEVDVFLKEEPALGDSBILKKQL
KQCRLLVSDIQTIPSLNSVNEGQKJKNABEPFASRLTELKELNTOQDMHCQQVY
ARKEALKGKLEKTVSLQKDLSEHWNTOAEYELERDEYKTPDELQKAVEEMKRAK
BEAQKEAKVLLTESVNSVIAQAPVVAQEAALKKELETITNTYNQWLCTRLNGCKTLE
EWAQWHELKANKWLNVEFEKLTENIPGAEEISEVLDLSLENLNRHSEDNF
NQIRILAQTLTDGVDDELINEELETNSRWRELHEEAVERKQLEQSIQSAQETERS
LHLIQESLTFIDKOLAAYIADKVDAAQPOEAQIQSDLTSHSIEISLEEMKHNQKKEA
AQVLSQIDVAQKQLQDVSMKFRLOKPAFANFELRLQKQWLENDEKMLPALETKSVE
QEVVQSLNHLVNLKSLSEKSEVEMVKTGRQIQVQKQNTENPKELDERVTALKLHY
NELGAKVTERKQKLEKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPNSLDSE
VAVGKATQKEIEKQVHLKSIIEVGEALKTVLGKKEIVEDKLSLNSNWIATVSRAE
EWNLLLEYQKHEMTFDQNVDTHTWIIQADTLTDESEKKKPPQKEDVLRKLKAEJND
IRPKVDSTRDQANLMAWGDHCKLVEPQISELNHRFAAISHRKTKGASIPKLELE
QFNSDIQKLEPLEABEIQQVNLKEEDFNKMDNEDNEGTVKELLQRGDNLQQRITDER
K3EEIKIQOQLQKHNALKDLRSORRKALEISHQWQYKQRAADDLLKCLDDIEKLL
ASLEPRDERKIKEDIRELQKKEELNANVRQAEGLSEDCAAAVEPTQIQLSKRWE
IESKFAQFRRLNFAQIHTVREETMTWMTEDMPEISVYVSTYLTETIHTVSQALLEVEQ
LLNAPDLCAKDFEDLFQKEESLKNIDSLQSSGRIDIHSHKTAALQSAATPVERVQL
QEALSQDLDFQWVKVWKYKDRQRFORSVEKRFHFVDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGQRTVVRTLNATGEEIIQSSKTDASILQELKLSNL
RWQEVCKQLSDRKRLEEQKNILSEFQDLNEFVLWLEHADNIAIPIEPKKEQQLKE
RALPEKQGEIEAQIKDLQLEKLEDEELQNLHLLWLSPIRNQLEIYNQPNQEGPFD
VQETEIAVQAKQPDVEEILSKGQHLKKEKPAQPVKRLKEDLSSEWKAVERNLLQELRA
KOPDLAPGLTTIGASPTQVTLVTPVVTKETAI SKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVKSQRVMVGDLINEMI IKQATWQDLEQRRPQLEELITAAQNLK
ELTDWLSLLDQVKSQRVMVGDLINEMI IKQATWQDLEQRRPQLEELITAAQNLK
NKTSNQEARITITDRIERIONOWDEQHLQNRQQLNEMKLDSTQWLEAKEEAEQVL
QOARAKLESKEGPTVDAIQKJITETQLAKOLQWQNTVDVANDLAKLLLRDYSAD
DTRKVMITENINASWRSIHKRVSERBAALTEHRLIQQFPFLDLEKFLAWLTAETTA
NVLDQATRERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLNDSQKILRSLEGS
DAVLLQNDNNMFKWSELKSLNIRSHLEASSDQWKRHLHLSQELLVWLQKDEL
SRQAPIGGDFPAVQKQNDVHRAFRKELTKPEVIMSTLETVRIFLTEQPLEGLEKLYQ
EPRELPEERAOQVTRLLRKQAEVNTWEKLNHSDAQWQKIDETLERLQLEQEAQD
ELDLKLRQAEVTKGSWPVGDLILDSLDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNRWLQVAVEDRVQRLHAEHDFGPASOHFLSTSVQGP
WERAI SPNVKVPYINHETQTTCDWHPKMTLEYQSLADLNNRFSAYRTAMKLRQLKA
LCDLILLSAACDALDQNLKQNDQPMIDLOI INCLTIIYDRLEQEHNLVNVPLCVD
MCLNWLNVYDTGRTRIRVLSFKTIGIISLCKAHEDKIRYLFKQVASTGFCDDQRL
GLLLHDSIQIPRQLGEVASFSGSNIEPSVRSFCQFANNKPEIEAALFLDWMRLPQSM
```

```
VWLPVLRHVAABETAKHQAQKCNICEKPIIGFRRYRSLKHFNYDICQSCFESGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKERTKRYFAKPRMGYLPVQTVLEGGDNMET
PVTLLINFPVDSAPASQSLSDDDTHSRIEHYAELAMENSNGSYLNDISIPNESID
DEHLLIQHYCOQLNQDSPLSQSPSPAQILISLESEERGERELILADLEENRNLAQAY
DRLKQHEHKLGLSPLPSPPEMPTSPQSPRDAELABAKLRQHKRLEARMQILEDH
NKQLESQHLRLQLLEQPOAEAKVNGTTVSSPSTQSRSDSSQPMLLRVVGSQTSDSM
GEEDLLSPPDQDTSTGLEEVMEQLNNSPPSSRGRNTPKPMREDTM"
ORIGIN
Query Match 63.2%; Score 1328.8; DB 9; Length 13957;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8598 GGTCCATTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 8657
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGG 881
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGG 8717
QY 882 ACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAA 941
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8718 ACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAA 8777
QY 942 CTAAGAAGCCTGTATATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8778 CTAAGAAGCCTGTATATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837
QY 1002 CTTTGAAGGACTAGAGAAAACCTTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAG 1061
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8838 CTTTGAAGGACTAGAGAAAACCTTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAG 8897
QY 1062 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8898 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 8957
QY 1122 AATTGAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 1181
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8958 AATTGAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 9017
QY 1182 AGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGTGAATCA 1241
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9018 AGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGTGAATCA 9077
QY 1242 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGA 1301
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9078 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGA 9137
QY 1302 AAGTCAAGGCACCTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACGCTGAGCCAGTCAATG 1361
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9138 AAGTCAAGGCACCTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACGCTGAGCCAGTCAATG 9197
QY 1362 ACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9198 ACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
QY 1422 TGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAAGC 1481
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9258 TGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAAGC 9317
QY 1482 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 1541
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9318 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 9377
QY 1542 TCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCTACTATATCAACACAG 1601
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9378 TCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCTACTATATCAACACAG 9437
QY 1602 AGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9438 AGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 9497
```


1662	ACCTGAATAATGTGAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA	1721
9498	ACCTGAATAATGTGAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA	9557
1722	AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAACACA	1781
9558	AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAACACA	9617
1782	ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAAATTTGTTTGACCACTA	1841
9618	ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAAATTTGTTTGACCACTA	9677
1842	TTTATGACCGCCTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATA	1901
9678	TTTATGACCGCCTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATA	9737
1902	TGTCCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC	1961
9738	TGTCCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC	9797
1962	TGTCCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	2021
9798	TGTCCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT	9857
2022	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC	2081
9858	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC	9917
2082	TTCTGCATGATTCTATCCAA	2101
9918	TTCTGCATGATTCTATCCAA	9937
RESULT 13		
AR220819		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
FEATURES		
source		
ORIGIN		
Query Match		
Best Local Similarity		
Matches 1333; Conservative		
0; Mismatches 7; Indels 1; Gaps 1;		
762	GGGTTCTTTTACAGACAGTTCTGACCAAGTGAAGCGTCTGACCTTTCTCTGCAAGAAC	821
8598	GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACCTTTCTCTGCAAGAAC	8657
822	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG	881
8658	TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG	8717
882	ACTTTCCAGCAGTTTCAAGACGAGAACGATGTATACATAGGGCCCTTCAAGAGGGAATTGAAA	941
8718	ACTTTCCAGCAGTTTCAAGACGAGAACGATGTATACATAGGGCCCTTCAAGAGGGAATTGAAA	8777
942	CTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	1001
8778	CTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	8837

QY	1002	CTTTGAAGGACTAGAGAACTCTACCAAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG	1061
Db	8838	CTTTGAAGGACTAGAGAACTCTACCAAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG	8897
QY	1062	CCCAGATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA	1121
Db	8898	CCCAGATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA	8957
QY	1122	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	1181
Db	8958	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	9017
QY	1182	AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA	1241
Db	9018	AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA	9077
QY	1242	AGGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	1301
Db	9078	AGGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA	9137
QY	1302	AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATG	1361
Db	9138	AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATG	9197
QY	1362	ACCTTGCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC	1421
Db	9198	ACCTTGCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC	9257
QY	1422	TGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGC	1481
Db	9258	TGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGC	9317
QY	1482	AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG	1541
Db	9318	AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG	9377
QY	1542	TCCAGGTTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACG	1601
Db	9378	TCCAGGTTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACG	9437
QY	1602	AGACTCAAAACAACCTTGTGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTTAGCTG	1661
Db	9438	AGACTCAAAACAACCTTGTGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTTAGCTG	9497
QY	1662	ACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA	1721
Db	9498	ACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA	9557
QY	1722	AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAAGCACA	1781
Db	9558	AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAAGCACA	9617
QY	1782	ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCACTA	1841
Db	9618	ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCACTA	9677
QY	1842	TTTATGACCGCCTGGAGCAAGACACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATA	1901
Db	9678	TTTATGACCGCCTGGAGCAAGACACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATA	9737
QY	1902	TGTCCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC	1961
Db	9738	TGTCCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC	9797
QY	1962	TGTCCTTTAAACTGSCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGAT	2021
Db	9798	TGTCCTTTAAACTGSCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGAT	9857
QY	2022	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCT-GGGCCTC	2080
Db	9858	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCGCTC	9917
QY	2081	CTTCTGCATGATTTCTATCCAA	2101

||||| 9918 CTTCTGCATGATTCTATCCAA 9938

SULT 14

0219

US Shortened dystrophin. 4402 bp DNA linear PAT 18-JUN-2001

FINITION

SSION E30219

SION E30219.1 GI:13017026

WORDS JP 1999318467-A/2.

URCE unidentified

ORGANISM unidentified

REFERENCE unclassified.

1 (bases 1 to 4402)

sinichi.T.

Shortened dystrophin

Patient: JP 1999318467-A 2 24-NOV-1999;

SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY

OS Unidentified

PN JP 1999318467-A/2

PD 24-NOV-1999

PF 08-MAY-1998 JP 1998142134

PR

PI SINICHI TAKEDA

PC C12N15/09,A61K48/00,C12N15/00

CC Strandedness: Both;

CC Topology: Linear;

FH Key Location/Qualifiers

FT source 1..4402 /organism='Unidentified'.

FT Location/Qualifiers

1..4402

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

NATURES

source

1

1095

61

1155

121

1215

181

1275

241

1335

301

1395

361

1455

421

1515

Query Match 62.7%; Score 1317; DB 6; Length 4402;

Best Local Similarity 81.6%; Pred. No. 0;

Matches 1714; Conservative 0; Mismatches 0; Indels 387; Gaps 1;

1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCATT 60

1095 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGAGGCCATT 1154

61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120

1155 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1214

121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180

1215 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1274

181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 240

1275 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA 1334

241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300

1335 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1394

301 TAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360

1395 TAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1454

361 AACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

1455 AACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1514

421 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 480

1515 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 1574

QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 540

Db 1575 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 1634

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600

Db 1635 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1694

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660

Db 1695 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1754

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720

Db 1755 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTA----- 1804

QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG 780

Db 1805 ----- 1804

QY 781 TTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 840

Db 1805 ----- 1804

QY 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCCAGCAGTTTCAGAA 900

Db 1805 ----- 1804

QY 901 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAATAACTAAAGAACCTGTATATCAT 960

Db 1805 ----- 1804

QY 961 GAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020

Db 1805 ----- 1804

QY 1021 ACTCTACCAGGAGCCACAGAGAGCTGCCTCTGTGAGGAGAGAGCCAGCAATGTCACTCGGCT 1080

Db 1805 ----- 1804

QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGC 1140

Db 1805 -----AGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGC 1847

QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGGCCAC 1200

Db 1848 TGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGGCCAC 1907

QY 1201 GGATGAGCTGGACCTCAAGCTCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260

Db 1908 GGATGAGCTGGACCTCAAGCTCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1967

QY 1261 GGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG 1320

Db 1968 GGGCGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG 2027

QY 1321 AGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC 1380

Db 2028 AGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC 2087

QY 1381 CACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1440

Db 2088 CACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 2147

QY 1441 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 1500

Db 2148 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 2207

QY 1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTGTCCAGGTCCTGGGAGAG 1560

Db 2208 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTGTCCAGGTCCTGGGAGAG 2267

y	1021	ACTCTAC	AGGAGCC	CAGAGAG	CTGCCTC	CTGAGG	AGAGCC	CAGAA	TGTCA	CTCGG	CT	1080
b	1990	ATGTTAT	CAAG									2000
y	1081	TCTAC	GAAAG	CAGGCT	GAGGAG	GTCAAT	ACTG	AGTGG	AAAAAT	TGAAC	CTG	1140
b	2001											2000
y	1141	TGACT	GGCAG	AGAAAA	ATAGAT	GAGACC	CTTG	AAAG	ACTCC	AGGA	ACTT	1200
b	2001											2000
y	1201	GGAT	GAGCT	GGAC	CTCAAG	CTG	CGCC	AAAG	CTG	AGGT	GATCA	1260
b	2001											2000
y	1261	GGGCG	ATCT	CTCAT	TGACT	CTCT	CC	AA	GATCA	CTCG	AGAA	1320
b	2001											2027
y	1321	AGAA	ATTG	CGCCT	CTG	AAAG	AGAA	CGT	GAGC	CACT	CAAT	1380
b	2028	AGAA	ATTG	CGCCT	CTG	AAAG	AGAA	CGT	GAGC	CACT	CAAT	2087
y	1381	CAC	TTTGG	GCAT	TTCAG	CTCT	CAC	CGT	ATA	AAC	CTC	1440
b	2088	CAC	TTTGG	GCAT	TTCAG	CTCT	CAC	CGT	ATA	AAC	CTC	2147
y	1441	ATG	GAAG	CTT	CTG	CAG	GTG	GG	CCGT	CG	AGG	1500
b	2148	ATG	GAAG	CTT	CTG	CAG	GTG	GG	CCGT	CG	AGG	2207
y	1501	GG	ACTT	TGG	TCC	AG	CAT	CTC	AG	CAC	GTCT	1560
b	2208	GG	ACTT	TGG	TCC	AG	CAT	CTC	AG	CAC	GTCT	2267
y	1561	AG	CCAT	CT	CG	CC	AA	CA	AA	AG	TGCC	1620
b	2268	AG	CCAT	CT	CG	CC	AA	CA	AA	AG	TGCC	2327
y	1621	GG	AC	AT	CCCC	AA	AT	G	A	C	A	1680
b	2328	GG	AC	AT	CCCC	AA	AT	G	A	C	A	2387
y	1681	CT	C	AG	CTT	A	G	A	CT	G	C	1740
b	2388	CT	C	AG	CTT	A	G	A	CT	G	C	2447
y	1741	CT	T	G	A	G	C	T	G	T	G	1800
b	2448	CT	T	G	A	G	C	T	G	T	G	2507
y	1801	G	C	C	A	T	G	A	T	A	T	1860
b	2508	G	C	C	A	T	G	A	T	A	T	2567
y	1861	A	G	A	C	A	A	A	T	T	G	1920
b	2568	A	G	A	C	A	A	A	T	T	G	2627
y	1921	G	A	A	T	T	G	T	A	T	A	1980
b	2628	G	A	A	T	T	G	T	A	T	A	2687
y	1981	C	A	T	T	C	C	T	G	T	A	2040
b	2688	C	A	T	T	C	C	T	G	T	A	2747
y	2041	A	A	G	T	T	C	A	C	A	G	2100
b	2748	A	A	G	T	T	C	A	C	A	G	2807

QY 2101 A 2101
Db 2808 A 2808

Search completed: April 5, 2004, 02:40:08
Job time : 5490.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

run on: April 4, 2004, 16:52:43 ; Search time 2274.66 Seconds
(without alignments)
17079.796 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	584	44.9	824	9	AL556247
2	583	44.8	777	14	CD653550
3	582.4	44.8	1098	13	BX365572
4	573	44.0	620	13	BQ640063

5	571.4	43.9	801	14	CB991394
6	537.8	41.3	770	12	BG719710
7	535	41.1	3753	11	AK081426
8	526.4	40.5	1298	11	AK087829
9	526.4	40.5	2135	11	AK013510
10	526.4	40.5	4437	11	AK036936
11	524.8	40.3	1384	11	AK075809
12	511	39.3	797	14	CB960722
13	508.2	39.1	2874	29	AY408546
14	502.8	38.6	652	10	BB629984
15	497	38.2	2874	29	AY408548
16	491.6	37.8	704	10	BB610411
17	489.8	37.6	508	14	CD701871
18	479.6	36.9	854	9	AI196693
19	470.8	36.2	717	14	CB527785
20	461.2	35.4	542	29	CG594848
21	457.2	35.1	513	29	CG605968
22	455.8	35.0	591	13	BX490860
23	453.8	34.9	508	29	CG606947
24	439	33.7	578	13	BX485574
25	436.6	33.6	665	13	BY742604
26	431.4	33.2	898	10	BF182065
27	427.8	32.9	843	14	CA988247
28	426.6	32.8	3051	11	BC036095
29	422.6	32.5	520	29	CG606443
30	420.2	32.3	512	14	CF162938
31	419.4	32.2	488	29	CG597965
32	416	32.0	493	29	CG604818
33	410.4	31.5	599	10	BB666688
34	407.2	31.3	554	14	CB613696
35	405.4	31.2	662	12	BJ626491
36	404.4	31.1	493	14	CA888041
37	397.4	30.5	495	14	CA894775
38	392.2	30.1	650	13	BY714491
39	389	29.9	794	13	BU424348
40	385.2	29.6	677	9	AL641565
41	383.4	29.5	483	14	CA893902
42	381.8	29.3	423	9	AA460476
43	380.4	29.2	622	9	AL796733
44	375.4	28.9	502	9	AL871560
45	365.4	28.1	655	14	CA558919

ALIGNMENTS

RESULT 1
AL556247
LOCUS AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 824)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK001CA09QPl.
Location/Qualifiers

AL556247 824 bp mrna linear EST 31-MAY-2003
AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK001YB17 5-PRIME, mRNA sequence.

source 1..824 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DK001YB17" /cell_type="HELA CELLS COT 25-NORMALIZED" /cell_line="HELA" /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 44.9%; Score 584; DB 9; Length 824; Best Local Similarity 100.0%; Pred. No. 6.1e-163; Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 718 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACCGTCTTT 777
Db 195 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACCGTCTTT 254

QY 778 ACCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
Db 255 ACCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 314

QY 838 GCAGAAGGCCCTTTGCTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 897
Db 315 GCAGAAGGCCCTTTGCTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 374

QY 898 GCACAACCTCAAGCABAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGAC 957
Db 375 GCACAACCTCAAGCABAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGAC 434

QY 958 CACTATTATGACCGCTGGAGCAAGAGCAACAATAATTGGTCAACGTCCTCTCTGCGT 1017
Db 435 CACTATTATGACCGCTGGAGCAAGAGCAACAATAATTGGTCAACGTCCTCTCTGCGT 494

QY 1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
Db 495 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 554

QY 1078 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAGTA 1137
Db 555 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAGTA 614

QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCGCAGGTGGG 1197
Db 615 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCGCAGGTGGG 674

QY 1198 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG 1257
Db 675 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG 734

QY 1258 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAATTG 1301
Db 735 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAATTG 778

RESULT 2 CD653550 777 bp mRNA linear EST 18-JUN-2003
LOCUS AGENCOURT_14553000 NIA Human H1 Embryonic Stem Cell cDNA Library
DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30426742 5', mRNA sequence.
ACCESSION CD653550
VERSION CD653550.1 GI:31892067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM513 row: a column: 23
High quality sequence stop: 715.
FEATURES Location/Qualifiers
1..777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30426742"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_hosts="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pgACTAGTCTAGATCGAGCGCGCCCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 44.8%; Score 583; DB 14; Length 777;
Best Local Similarity 99.8%; Pred. No. 1.2e-162;
Matches 583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 718 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACCGTCTTT 777
Db 137 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACCGTCTTT 196

QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGTTATAGGACTGCCATGAAACTCCGAAGACT 837
Db 197 AGCTGACCTGAATAATGTCAGATTCTCAGTTATAGGACTGCCATGAAACTCCGAAGACT 256

QY 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 897


```
|||||
db 257 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 316
|||||
iy 898 GCACAACCTCAAGCAAAATGACGACCCCATGATATCCTGCAGATTATTAATTGTTTGAC 957
|||||
ib 317 GCACAACCTCAAGCAAAATGACGACCCCATGATATCCTGCAGATTATTAATTGTTTGAC 376
|||||
y 958 CACTATTATGACCGCCTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGGGT 1017
|||||
b 377 CACTATTATGACCGCCTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGGGT 436
|||||
y 1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
|||||
b 437 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 496
|||||
y 1078 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
|||||
b 497 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 556
|||||
y 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGG 1197
|||||
b 557 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGG 616
|||||
y 1198 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
|||||
b 617 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 676
|||||
y 1258 GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
|||||
b 677 GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCNATTG 720
|||||

RESULT 3
X365572 1098 bp mRNA linear EST 05-MAY-2003
OCUS BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
EFINITION CDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
CCESSION BX365572
ERSION BX365572
EYWORDS EST.
OURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1098)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AK001ZB11QP1.
Location/Qualifiers
1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

EATURES
source
1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

RIGIN

Query Match 44.8%; Score 582.4; DB 13; Length 1098;
```

```
Best Local Similarity 99.8%; Pred. No. 2.3e-162;
Matches 583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 718 CCACGAGACTCAAAACAACCTTCTCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 777
|||||
Db 204 CCACGAGACTCAAAACAACCTTCTCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 263
|||||
QY 778 AGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
|||||
Db 264 AGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 323
|||||
QY 838 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCCTTGACCA 897
|||||
Db 324 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCCTTGACCA 383
|||||
QY 898 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 957
|||||
Db 384 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 443
|||||
QY 958 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 1017
|||||
Db 444 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 503
|||||
QY 1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
|||||
Db 504 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 563
|||||
QY 1078 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTA 1137
|||||
Db 564 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTA 623
|||||
QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGG 1197
|||||
Db 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGG 683
|||||
QY 1198 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
|||||
Db 684 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 743
|||||
QY 1258 GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
|||||
Db 744 GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 787
|||||

RESULT 4
BQ640063 620 bp mRNA linear EST 15-JUL-2002
LOCUS he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
DEFINITION Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION BQ640063
VERSION BQ640063.1 GI:21764522
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 620)
Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), 196-204 (2002)
22103461
12107411
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
```

Seq primer: M13RP1 reverse primer (ABI).

FEATURES
source
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor
[5'-pGACTAGTCTAGATCGGCGCGCCGCTT]5-3'. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 44.0%; Score 573; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 9.7e-160;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 729 AAACAACCTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGA 788
Db 1 AAACAACCTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGA 60
QY 789 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCC 848
Db 61 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCC 120
QY 849 TTTGCTTGGATCTCTTGGCTGTGTCAGCTGCATGTGATGCCCTTGGACGACACACCTCA 908
Db 121 TTTGCTTGGATCTCTTGGCTGTGTCAGCTGCATGTGATGCCCTTGGACGACACACCTCA 180
QY 909 AGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTTGTTGACCACTATTATG 968
Db 181 AGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTTGTTGACCACTATTATG 240
QY 969 ACCGCTGGAGCAAGAGCAACAATTTGTTCAACGTCCCTCTCGTGGATATGTGTC 1028
Db 241 ACCGCTGGAGCAAGAGCAACAATTTGTTCAACGTCCCTCTCGTGGATATGTGTC 300
QY 1029 TGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCTGCTTT 1088
Db 301 TGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCTGCTTT 360
QY 1089 TTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTT 1148
Db 361 TTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTT 420
QY 1149 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTGTC 1208
Db 421 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTGTC 480
QY 1209 ATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAAGTAA 1268
Db 481 ATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAAGTAA 540
QY 1269 TTGAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 541 TTGAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 573

RESULT 5

CB991394
LOCUS
DEFINITION
AGENCOURT_13627932 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
CB991394
ACCESSION
CB991394.1 GI:30285818
VERSION
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 801)
REFERENCE
1 NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.

FEATURES

source
Location/Qualifiers
1..801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30336570"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 43.9%; Score 571.4; DB 14; Length 801;
Best Local Similarity 99.7%; Pred. No. 3.5e-159;
Matches 583; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 718 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTT 777
Db 140 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTT 199
QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
Db 200 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 259
QY 838 GCAGAGGCGCTTTGTTGATCTCTTGGAGCTGTGAGCTGCATGTGATGCCTTGACCA 897
Db 260 GCAGAGGCGCTTTGTTGATCTCTTGGAGCTGTGAGCTGCATGTGATGCCTTGACCA 319
QY 898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTGTTGAC 957
Db 320 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTGTTGAC 379
QY 958 CACTATTATGACCGCCTTGAGCAAGAGCAACAATTTGGTCAACGCTCCTCTCTGCGT 1017
Db 380 CACTATTATGACCGCCTTGAGCAAGAGCAACAATTTGGTCAACGCTCCTCTCTGCGT 439

1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1077
440 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 499
1078 TGTCCTGCTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
500 TGTCCTGCTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 559
1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGGCTGGG 1197
560 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGGCTGGG 619
1198 CCTCCTTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTGTGGTGAAGTTGCATCCTTTGG 1257
620 CCTCCTTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTGTGGTGAAGTTGCATCCTTTGG 679
1258 GGGCAGTAACATTGAGCCCAAGTGTCC-GGAGCTGCTTCCAATTG 1301
680 GGGCAGTAACATTGAGCCCAAGTGTCCGGGAGCTGCTTCCAATTG 724

RESULT 6
719710
LOCUS 770 bp mRNA linear EST 08-MAY-2001
DEFINITION 602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5',
mRNA sequence.
719710
719710.1 GI:13998897
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 770)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10731 row: 1 column: 08
High quality sequence stop: 767.
Location/Qualifiers
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822807"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

FEATURES
Source
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822807"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

Query Match 41.3%; Score 537.8; DB 12; Length 770;
Best Local Similarity 97.4%; Pred. No. 4e-149;
Matches 568; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

IGIN

718 CCACGAGACTCAAAACAACCTTGTCTGGACCACATCCCAAAATGACAGAGCTCTACCACTCTT 777
150 CCACGAGACTCAAAACAACCTTGTCTGGACCACATCCCAAAATGACAGAGCTCTACCACTCTT 209
778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
210 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 269
838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCA 897
270 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCA 329
898 GCACAA--CCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTG 955
330 GCACAATCCTCAAGTCAGAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTG 389
956 ACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGC 1015
390 ACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGC 449
1016 GTGGATATGTGCTGAACCTGGCTGGTGAATGTTATGATACGGGACGAACAGGAGGATC 1075
450 GTGGATATGTGCTGAACCTGGCTGGTGAATGTTATGATACGGGACGAACAGGAGGATC 509
1076 CGTGTCTCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAG 1135
510 CGTGTCTCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAG 569
1136 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCGAGCTG 1195
570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGA-TTTGTGACCAAGCGAGCTG 628
1196 GGCCTCCTTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTGGGTGAAGTTGCATCCTTT 1255
629 GGCCTCCTTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTGGGTGAAGTTGCATCCTTT 688
1256 GGGGGCAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAT 1298
689 GGGGGCAGTAACATAGAGCCCAAGTGTCCGGAGCTGTTCCAAT 731

RESULT 7
AK081426
LOCUS 3753 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
library, clone:Cl30016K19 product:dystrophin related protein 2,
full insert sequence.
AK081426
AK081426.1 GI:26349154
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Y 794 GTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGC 853
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1372 ATTAAGTTCTCAGCTTACCGCACTGCCATGAAGCTCCGCAGAGTCCAGAAGGCCCTGCGC 1431
Y 854 TTGGATCTCTTGACCTGTTCAGCTGCATGTGTGATGCCCTTGGACCAGCACAACTCAAGCAA 913
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1432 CTGGATCTGGTAACCTTTAACTACAGCTCTGGAGATCTTCAATGAGCATGACTTGCAGGCC 1491
Y 914 AATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCACTATTATTGACCCGC 973
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1492 ACTGAACATGTGATGGATGTGGTGGAGGTCAATCACTGCTGAATGCCTTGATGAACGA 1551
Y 974 CTGGAGCAAGAGCACAACAATTGTTGCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC 1033
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1552 CTGGAGGAGGAAGAGGCATCCTGGTCAATGTGCCACTGTGTAGACATGAGCCTCAAC 1611
Y 1034 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTGTCTCTTTTAAA 1093
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1612 TGGCTCCTCAATGTTTGTATAGTGTGCGAGTGGAAAGATGCGAGCATTTGTCCTTAAAG 1671
Y 1094 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAGATACCTTTTCAAG 1153
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1672 ACTGGCATCGCATGCTGTGTGGCACGGAAGTGAAGAGAAAAAATTCAGTATCTCTTCAGC 1731
Y 1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGAGGTGGGCTCTCTTCTGTCATGAT 1213
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1732 CAAGTAGCCAATTCAGGCAGCCAGTGTGATCAACGCCATCTCGGTGCCCTGCTTCATGAA 1791
Y 1214 TCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCTCTTTGGGGCAGTAACATTGAG 1273
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1792 GCCATTCAAGTGCCCGCTCAGCTGGTGAAGTGGCAGCATTTGGGGCAGCAATGTGGAG 1851
Y 1274 CCAAGTGTCCGGAGCTGCTTCCAAATTT 1300
b | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1852 CCCAGTGTCCGTAGCTGCTTTCCGCTTT 1878

AK087829 1298 bp mRNA linear HTC 20-SEP-2003
Mus musculus 2 days pregnant adult female ovary cdna, RIKEN
full-length enriched library, clone:E330026B12 product:dystrophin,
muscular dystrophy, full insert sequence.
AK087829
AK087829.1 GI:26352777
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cdna libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazana,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cdna collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 50,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1298)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cdna library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. 1298
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330026B12"
/db_xref="MGI:2426619"
/db_xref="taxon:10090"
/clone="E330026B12"
/sex="female"
/tissue type="ovary"
/clone_lib="RIKEN full-length enriched mouse cdna library"
/dev_stage="2 days pregnant adult"
120. 1142
/note="unnamed protein product; dystrophin, muscular
dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN,
100% match=999)
putative"
/codon_start=1
/protein_id="BAC40019.1"
/db_xref="GI:26352778"
/translation="MREHLKGHETQTTCWDHPKMTELYQSLADLNVRFSAYRTAMKL
RRLOKALCLDLLSACDALDQHNLKONDPMDILOINCLTTIYDRLEQEHNNLVN
VPLCDVMDCLNLLNVYDTGRTGIRVLSTFKGIIISLCAHLEDKYRLFQVASTGF
CDORRLGLLLHDSIQIPQLGEVASFGSNTEPSVRSFQFANNKPEFAALFLDWMR
LEFQSMVWLPVLHRVAAETAQHQAACNIKCEPIIGFRRYSLKHFNYDICQSCFFSG
RVAKGHKMHYPMVEYCTPTTSGEDVRDFAVLKNKFKRTKRYFAKHPRMGVLPVQTVLE
GDNNET"

FEATURES
source

CDS


```
misc_feature 1..2135
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)"
RIGIN
Query Match 40.5%; Score 526.4; DB 11; Length 2135;
Best Local Similarity 93.8%; Pred. No. 2.2e-145;
Matches 548; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
718 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTTACCAGTCTTT 777
|||||
144 CCACGAGACCCAAACCACTTGTTGGGACCACCCCAAATGACAGAGCTTACCAGTCTTT 203
|||||
778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACT 837
|||||
204 AGCTGACCTGAATAATGTCAGGTTCTCCGCGTATAGGACTGCCATGAAGTCAGAAGGCT 263
|||||
838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA 897
|||||
264 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA 323
|||||
898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATGTTTGAC 957
|||||
324 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATAATTAAGTGTGGAC 383
|||||
958 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 1017
|||||
384 TACAATTATGATCGTCTGGAGCAAGAGCACAAATCTGGTCAATGTCCTCTCTGTGT 443
|||||
1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1077
|||||
444 GGATATGTGCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCG 503
|||||
1078 TGTCCTGTCTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
|||||
504 TGTCCTGTCTTTAAACTGGCATATTTCTCTGTGTAAAGCACACTTGGAGACAAGTA 563
|||||
1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCACACAGGATTTGTGACCAGCGCAGGCTGGG 1197
|||||
564 CAGATACCTTTCAAGCAAGTGGCAAGTTCACACTGGCTTTGTGACCAGCGTAGGCTGGG 623
|||||
1198 CCTCCTTTCGATGATCTCTATCCAAATCCAGACAGTGGGTGAAGTTGATCCTTTGG 1257
|||||
624 TCTTCTTTCGATGATCTCTATTCAAATCCCAAGACAGTGGGTGAAGTTGCTCCTTTGG 683
|||||
1258 GGGCAGTAACATTGAGCCCAAGTGTCGGAGCTGCTTCCAATTG 1301
|||||
684 GGGCAGTAACATTGAGCCGAGTGTCAGGAGCTGCTTCCAATTG 727
|||||
AK036936 4437 bp mRNA linear HTC 19-SEP-2003
Mus musculus adult female vagina cDNA, RIKEN full-length enriched
library, clone:9930028B14 product:dystrophin, muscular dystrophy,
full insert sequence.
AK036936
AK036936.1 GI:26085504
HTC; CAP trapper.
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
3REFERENCE
1
AUTHORS
Carninci,P. and Hayashizaki,Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2
3REFERENCE
2
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
```

```
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yokeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4437)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
1..4437
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:9930028B14"
/db_xref="MGI:2401399"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
29..1968
misc_feature
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)"
```


/note="unnamed protein product; dystrophin, muscular
dystrophy (MGD|GI:94909, GB|NM_007868, evidence: BLASTN,
100%, match=999)
putative"
/codon start=1
/protein_id="BAC35977.1"
/db_xref="GI:26344656"
/translation="MREHLKGHETQTTCDWHPKWTELYQSLADLNNVRFSAYRTAMKL
RRLQKALCLDLLSAACDALDQHNLKQNDQPMILQIINCLTTIYDRLEQEHNNLVN
VPLCVDMCLNLLNVYDTGRIRVLSFKTGISLCKAHLETKYLFKQVASSTGF
CDQRRLGLLHDSIQIPRLGEVAFSGSNIEPSVRSFCQFANNKPEIEAALFLDWMR
LEPQSMVWLPLVLRVAAATAKHQAKCNICEPIIGFPRYSLKHFNYDICQSCFFSG
RVAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFKTKRYFAKHPRMGYLPVQTVLE
GDNMET"

RIGIN

Query Match	40.3%;	Score 524.8;	DB 11;	Length 1384;
Best Local Similarity	93.7%;	Pred. No. 4.8e-145;		
Matches 547;	Conservative 0;	Mismatches 37;	Indels 0;	Gaps 0;
718	CCACGAGACTCAAACTTCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT	777		
140	CCACGAGACCCAAACCACTTGTGGGACCAACCCCAAAATGACAGAGCTCTACCACTCTTT	199		
778	AGCTGACCTGAATAATGTCTGATCTCTGAGCTTATAGGACTGCCATGAACTCCGAAGACT	837		
200	AGCTGACCTGAATAATGTCTGAGTTCTCCGGGTATAGGACTGCCATGAAGCTCAGAAGCT	259		
838	GCAGAAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA	897		
260	CCAGAAAGGCCCTTTGCTTGGATCTCTGAGCCTGTCCGGTGCATGTGATGCCCTTGGACCA	319		
898	GCACAACTCAAGCAAAATGACCGCCCATGGATATCTTCGAGATTATTAAATGTTTGAC	957		
320	GCACAACTCAAGCAAAATGACCGCCCATGGATATCTTCGAGATAATTAATGTTTGAC	379		
958	CACATATTTATGACCGCCTGGAGCAAGACACAACAATTTGGTCAACGTCCTCTCTGCGT	1017		
380	TACAATTTATGATCGTCTGGAGCAAGACACAACAATCTGGTCAATGTCCCTCTCTGTGT	439		
1018	GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG	499		
1078	TGTCCTGTCTTTTAAACTGGCAAGTTCCTGTGTAAGCACATTTGGAAGACAAGTA	1137		
500	TGTCCTGTCTTTTAAACTGGCAAGTTCCTGTGTAAGCACACTTTGGAAGACAAGTA	559		
1138	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCGAGGCTGGG	1197		
560	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCGAGGCTGGG	619		
1198	CCTCCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTGGGTGAAGTTCATCCTTTGG	1257		
620	TCTTCTTCTGCATGATTTCTATTCAAATCCCAAGACAGTGGGTGAAGTTCCTTTGG	679		
1258	GGGCAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCCAATTG	1301		
680	GGGCAGTAACATTGAGCCGAGTGTCCGAGCTGCTTCCCAATTG	723		

RESULT 12
3960722
JCUS
EFINITION
AGENCOURT 13761995 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
CB960722
CB960722
CB960722.1 GI:30216839
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 797)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers
1. 797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site:1:
all-XhoI; Site:2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

FEATURES
source

ORIGIN

Query Match	39.3%;	Score 511;	DB 14;	Length 797;
Best Local Similarity	99.8%;	Pred. No. 4.3e-141;		
Matches 522;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
QY	718	CCACGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT	777	
Db	168	CCACGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT	227	
QY	778	AGCTGACCTGAATAATGTCTGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT	837	
Db	228	AGCTGACCTGAATAATGTCTGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT	287	
QY	838	GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCCTTGGACCA	897	
Db	288	GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCCTTGGACCA	347	
QY	898	GCACAACTCAAGCAAAATGACCGCCCATGGATATCTTCGAGATTATTAAATGTTTGAC	957	
Db	348	GCACAACTCAAGCAAAATGACCGCCCATGGATATCTTCGAGATTATTAAATGTTTGAC	407	
QY	958	CACATATTTATGACCGCCTGGAGCAAGACACAACAATTTGGTCAACGTCCTCTCTGCGT	1017	
Db	408	CACATATTTATGACCGCCTGGAGCAAGACACAACAATTTGGTCAACGTCCTCTCTGCGT	467	
QY	1018	GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG	1077	
Db	468	GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG	527	
QY	1078	TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA	1137	
Db	528	TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA	587	
QY	1138	CAGATACCTTTTCAAGCAAGTGGCAAGTTCGAAGTTCGAAGTTCGAAGTTCGAAGTTCGG	1197	
Db	588	CAGATACCTTTTCAAGCAAGTGGCAAGTTCGAAGTTCGAAGTTCGAAGTTCGAAGTTCGG	647	
QY	1198	-CCTCCTTCTGCATGATTTCTATCCAAATTCGAAGTTCGAAGTTCGG	1239	

URCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Muridae; Murinae; Mus.
1 (bases 1 to 652)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
Location/Qualifiers
1..652
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult female vagina"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN
Query Match 38.6%; Score 502.8; DB 10; Length 652;
Best Local Similarity 92.6%; Pred. No. 1.1e-138;
Matches 539; Conservative 0; Mismatches 42; Indels 1; Gaps 1;
QY 718 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTTACAGTCTTT 777
Db 45 CCACGAGACCCAAACCACTTGTTGGGACCAACCCCAAAATGACAGAGCTTACAGTCTTT 104
QY 778 AGTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 837
Db 105 AGTGACCTGAATAATGTCAGGTTCTCCGCGATAGGACTGCCATGAGCTCAGAAGGCT 164
QY 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCA 897
Db 165 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCA 224
QY 898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTTGAC 957
Db 225 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTTGAC 284
QY 958 CACTATTATGACCGCTTGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCGT 1017
Db 285 TACAATTATGATCGTCTGGAGCAAGAGACACAACAATCTGGTCAATGTCCTCTCTGTGT 344
QY 1018 GGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1077
Db 345 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCG 404
QY 1078 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
Db 405 TGTCTGTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCACACTTGGAAACAAGTA 464
QY 1138 CAGATACCTTTTCAAGCAAGTGCGCAAGTTCAACAGGATTTTGTGACCAGCGAGGCTGGG 1197
Db 465 CAGATACCTTTTCAAGCAAGTGCGCAAGTTCAACTGGCTTTTGTGACCAGCGAGGCTGGG 524
QY 1198 CCTCCTTCTGCATGATTCTATCAAAATTCGAAGACAGTGGGTGAAGTTGCATCCTTTGG 1257
Db 525 TCTTCTTCTGCATGATTCTATCAAAATCCCAAGACACGTTGGGTGAAGTTGCTTCTTTC 584
QY 1258 GGGCAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCAATT 1299
Db 585 GGGCA-TAACATTGAGCCGAGTGTGAGGAGCTGTTTCAATT 625

RESULT 15
AY408548
LOCUS
DEFINITION Mus musculus DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY408548
VERSION AY408548.1 GI:39764519
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2874)
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2874)

Job time : 2280 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:58:13 ; Search time 97.0414 Seconds
(without alignments)
11443.127 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaca 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182.6	59.1	5952	4	US-09-687-875A-1
2	1182.6	59.1	13977	4	US-09-484-970B-60
C 3	1004	50.2	19307	3	US-08-836-022A-10
C 4	1004	50.2	19307	3	US-09-427-048A-10
5	397	19.8	6045	4	US-09-091-501B-7
6	397	19.8	10320	4	US-09-091-501B-9
7	311.8	15.6	3915	4	US-09-976-594-93
8	79.4	4.0	200	4	US-09-091-501B-5
9	78.6	3.9	200	4	US-09-091-501B-4
10	78.6	3.9	200	4	US-09-091-501B-6
C 11	76.6	3.8	7218	1	US-08-232-463-14
12	63.6	3.2	238	4	US-09-687-875A-13
C 13	46.2	2.3	505	4	US-09-621-976-15639
C 14	44	2.2	1230025	4	US-09-198-452A-1
15	43.4	2.2	1179	4	US-09-107-532A-1186
16	42.8	2.1	1690	4	US-09-620-312D-69
17	42.8	2.1	7812	3	US-09-368-590-1
18	40.4	2.0	832	4	US-09-621-976-2813
19	40.4	2.0	2223	1	US-08-257-073-4
20	39.2	2.0	16995	4	US-08-961-527-82
21	38.6	1.9	1751	4	US-09-620-312D-847
22	38.6	1.9	1995	1	US-08-425-069-3
23	38.6	1.9	1995	2	US-08-317-844B-3
24	38.4	1.9	7672	4	US-09-220-132-24
25	38.2	1.9	1131	6	5180810-3
26	38.2	1.9	1784	6	5180810-2
C 27	38	1.9	1394	4	US-09-247-155-76
					Sequence 76, Appl

28	36.8	1.8	1886	6	5210183-1	Patent No. 5210183
29	36.6	1.8	1845	4	US-08-887-534A-22	Sequence 22, Appl
30	36.6	1.8	1845	4	US-09-527-431-22	Sequence 22, Appl
C 31	36.6	1.8	7075	4	US-08-956-171E-263	Sequence 263, App
32	36.2	1.8	289	3	US-09-007-005-17	Sequence 17, Appl
33	36.2	1.8	289	3	US-09-244-796-17	Sequence 17, Appl
34	36.2	1.8	2447	2	US-09-014-969-14	Sequence 14, Appl
C 35	36.2	1.8	11049	4	US-10-204-708-23	Sequence 23, Appl
36	36.2	1.8	168575	4	US-09-426-290-1	Sequence 1, Appl
C 37	36	1.8	724	4	US-08-956-171E-832	Sequence 832, App
38	36	1.8	2873	4	US-08-630-915A-193	Sequence 193, App
39	35.8	1.8	790	3	US-09-461-474-13	Sequence 13, Appl
40	35.6	1.8	4868	1	US-08-139-937-12	Sequence 12, Appl
41	35.6	1.8	4868	5	PCT-US93-11310-12	Sequence 12, Appl
42	35.6	1.8	5934	4	US-09-418-710-2	Sequence 2, Appl
43	35.6	1.8	8257	4	US-09-595-684B-30	Sequence 30, Appl
44	35.6	1.8	8789	1	US-08-328-254-5	Sequence 5, Appl
45	35.6	1.8	10136	1	US-08-353-700-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED P
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match	59.1%;	Score 1182.6;	DB 4;	Length 5952;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1185;	Conservative	0;	Mismatches	4;
			Indels	0;
			Gaps	0;
Qy	813	GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCTG	872	
Db	2946	GGAAGAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCTG	3005	
Qy	873	GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCT	932	
Db	3006	GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCT	3065	
Qy	933	CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA	992	
Db	3066	CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA	3125	
Qy	993	AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAG	1052	
Db	3126	AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAG	3185	
Qy	1053	ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGATAACATGAAC	1112	

Db 3186 ATCCCTGGAAGGTTCCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCTT 3245
Qy 1113 CAAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
Db 3246 CAAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 3305
Qy 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 1232
Db 3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 3365
Qy 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAACA 1292
Db 3366 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAACA 3425
Qy 1293 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
Db 3426 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 3485
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAAACT 1412
Db 3486 TACTCTTGAGACTGTACGAATATTTCTTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAAACT 3545
Qy 1413 CTACCAAGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGCCCGCAGAAATGTCACTCGGCTTCT 1472
Db 3546 CTACCAAGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGCCCGCAGAAATGTCACTCGGCTTCT 3605
Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 1532
Db 3606 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 3665
Qy 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
Db 3666 CTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACCTCAAGAGGCCACGGA 3725
Qy 1593 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 3726 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 3785
Qy 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 3786 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3845
Qy 1713 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTTACCAC 1772
Db 3846 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTTACCAC 3905
Qy 1773 TTTGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832
Db 3906 TTTGGCATTTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 3965
Qy 1833 GAAGCTTCTGCAGGTGCGCTGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGA 1892
Db 3966 GAAGCTTCTGCAGGTGCGCTGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGA 4025
Qy 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGGC 1952
Db 4026 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGGC 4085
Qy 1953 CATCTGCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAACA 2001
Db 4086 CATCTGCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAACA 4134

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 59.1%; Score 1182.6; DB 4; Length 13977;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
Db 8260 GGAAGAAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 8319
Qy 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db 8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCT 8379
Qy 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA 992
Db 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA 8439
Qy 993 AATTGAAGCTCACACAGATGTTTATCAAACTTGGATGAAAAAGCAAGCAAAATCCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCAAACTTGGATGAAAAAGCAAGCAAAATCCTGAG 8499
Qy 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCT 1112
Db 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCT 8559
Qy 1113 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTC 1172
Db 8560 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTC 8619
Qy 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAA 8679
Qy 1233 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 1292
Db 8680 AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 8739
Qy 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352
Db 8740 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 8799
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8859
Qy 1413 CTACCAAGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 8860 CTACCAAGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 8919
Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 8979
Qy 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
Db 8980 CTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA 9039
Qy 1593 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGG 1652

Db 9040 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGATCAAGGGATCTTGGCAGCCCGTGGG 9099
Qy 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
Qy 1713 AATTGGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCGCCAGCTTACCAC 1772
Db 9160 AATTGGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCGCCAGCTTACCAC 9219
Qy 1773 TTTGGGCATTGACTCTCAACCGTATAAACCCTCAGCAGCTCTGGAAGACCTGAACACAGATG 1832
Db 9220 TTTGGGCATTGACTCTCAACCGTATAAACCCTCAGCAGCTCTGGAAGACCTGAACACAGATG 9279
Qy 1833 GAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCTCAGGGA 1892
Db 9280 GAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCTCAGGGA 9339
Qy 1893 CTTTGGTCCAGCATCTCAGCATTCTTCTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGGC 1952
Db 9340 CTTTGGTCCAGCATCTCAGCATTCTTCTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGGC 9399
Qy 1953 CATCTCGCCAAACAAAGTGGCCTACTATATCAACACGAGACTCAACA 2001
Db 9400 CATCTCGCCAAACAAAGTGGCCTACTATATCAACACGAGACTCAACA 9448

RESULT 3

US-08-836-022A-10/C
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA

US-08-836-022A-10
Query Match 50.2%; Score 1004; DB 3; Length 19307;
Best Local Similarity 90.3%; Pred. No. 1.le-302;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTG 872
Db 6434 GGAAGAAACTCATAGATTACTGCAAGCAGTTCCCTCTGGACCTGGAGAAGTTTCTTGCCCTG 6375
Qy 873 GCTTACAGAAAGCTGAAACCAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGAAAAGGCT 932
Db 6374 GATTACGGAAGCAGAAACCAACTGCCAATGTCTCTACAGGACGCTTCCCGTAAGGAGAAGCT 6315
Qy 933 CCTAGAAAGACTCCAAGGGAGTAAAGAGAGCTGATGAAAACAATGGCAAGACCTTCAAGGTGA 992
Db 6314 CCTAGAAAGACTCCAAGGGAGTCAAGAGAGCTGATGAAAACAATGGCAAGATCTTCAAGGAGA 6255
Qy 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCTCTGAG 1052
Db 6254 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCTCTGAG 6195
Qy 1053 ATCCCTGGAAGGTTCCGATGATGCAAGTCTCTCAACATTTAGGTCCTTGGATAACATGAATTT 1112
Db 6194 ATCCCTGGAAGGTTCCGATGATGCAAGTCTCTCAACATTTAGGTCCTTGGATAACATGAATTT 6135
Qy 1113 CAAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATTTAGGTCCTTGGATAACATGAATTT 1172
Db 6134 CAAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATTTAGGTCCTTGGATAACATGAATTT 6075
Qy 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 6074 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 6015
Qy 1233 AGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCACTTCCAGCAGTTTCAGAGCA 1292
Db 6014 AGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCACTTCCAGCAGTTTCAGAGCA 5955
Qy 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTTAAAGAACCTGTATCATGAG 1352
Db 5954 GAATGATATACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTTAAAGAACCTGTATCATGAG 5895
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
Db 5894 TACTCTTGAGACTGTGAGAAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 5835
Qy 1413 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 5834 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 5775
Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAATTAAGAACTTCAAGAGGCGCAGCA 1532
Db 5774 ACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAATTAAGAACTTCAAGAGGCGCAGCA 5715
Qy 1533 CTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCGCAGCA 1592
Db 5714 TTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCGCAGCA 5655
Qy 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 5654 TGAACCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 5595
Qy 1653 CGATCTCCTCATGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 5594 GGATCTCCTCATGACTCTCTCCAAGATCACCTTCGAGAAAGTCAAGGCACCTTCGAGGAGA 5535
Qy 1713 AATTGGCGCTCTGAAAAGAGAACCGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCAC 1772
Db 5534 AATTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAATGACCTTGCACATCAGCTGACCAC 5475
Qy 1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCATGATG 1832
Db 5474 ACTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTTTTGAAGATCTGAATACCAGATG 5415

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7

LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open

US-09-091-501B-7

Query Match 19.8%; Score 397; DB 4; Length 6045;
Best Local Similarity 59.3%; Pred. No. 4.2e-113;
Matches 698; Conservative 0; Mismatches 470; Indels 9; Gaps 1;

2Y	833	TGCAACAGTTCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAACAA	892
Db	3069	TGCAGGCCCTCTGCAGAGATCTGGAAAACCTTCTGAAGTGGATCCAAGAGCAGAGACCA	3128
2Y	893	CTGCCAATGTCTACAGGATGTCTACCCGTAAAGAAAGGCTCTAGAAAGACTCCAAGGGAG	952
Db	3129	CAGTGAATGTGTTGTGGATGCCCTCTCATCGGGAGAAATGCTCTCTCAGGATAGTATCTTGG	3188
2Y	953	TAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATG	1012
Db	3189	CCAGGGAACCTCAAACAGCAGATGCAGACATCCAGGCAGAAATTGATGCCCAACATGACA	3248
2Y	1013	TTTATCACAACCTGGATGAAAACAGCCAAAATACTCTGAGATCCCTGGAAGGTTCCGATG	1072
Db	3249	TATTTAAAGCATGTACGGAACAGGCAGAAAGATGGTAAAGCTTTGGGAAATCTGAAG	3308
2Y	1073	ATGCAGTCTGTTCACAAAGACGTTTGGATAACATGAATCAAGTGGAGTGAATTCGGA	1132
Db	3309	AGGCTACTATGCTTCAACATCGACTGGATGATGATGAACCAAGATGGAATGACTTAAAG	3368
2Y	1133	AAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGAAGCGTCTGC	1192
Db	3369	CAAAATCTGCTAGCATCAGGGCCCAATTTGGAGCCAGCGCTGAGAAAGTGAACAGGTTGC	3428
2Y	1193	ACCTTCTCTGCAAGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCGGC	1252
Db	3429	TGATGTCCTTAGAAGAACTGATCAAAATGGTGAATATGAAGATGAAGAGCTTAAGAAAC	3488
2Y	1253	AGGCACCTATTGGAGCGCACTTTCCAGCACTTCAGAAAGCAGAACGATGTACATAGGCGCT	1312
Db	3489	AAATGCCTATTGGAGAGATGTTCCAGCCTTACAGCTCCAGTATGACCATTTGTAAGGCC	3548

QY	1313	TCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTTGTGAGACTGTACGAA	1372
Db	3549	TGAGACGGGAGTTAAAGGAGAAAGAAATATCTGTCTCTGAATGCTGTGACCCAGGCCGAG	3608
QY	1373	TATTTCTGACAGAGCAGCCTTT-----GGAAGGACTAGAGAAACTCTACCGAGGAGC	1423
Db	3609	TTTTCTTGGCTGATCAGCCCAATTGAGGCCCTGAAGAGCCCAAGAAACCTACAATCAA	3668
QY	1424	CCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGG	1483
Db	3669	AAACAGAAATTAATCTCTGAGGAGAGAGCCCAAAAGATTGCCAAAGCCATGCGCAACAGT	3728
QY	1484	CTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCATCTCGCTGACTGGCAGAGAA	1543
Db	3729	CTTCTGAAGTCAAGAAAAAATGGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGC	3788
QY	1544	AAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACC	1603
Db	3789	AAGTGGACAAGGCATTGGAGAAACTCAGAGACCTGCAGGGAGTATGGATGACCTGGAGC	3848
QY	1604	TCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCTGGCGGATCTCCTCA	1663
Db	3849	CTGACATGAAGGAGCAGAGTCCGTGCGGAATGGCTGGAAGCCCTGGGAGACTTACTCA	3908
QY	1664	TTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTGGCGCTC	1723
Db	3909	TTGACTCGCTGCAGATCACATTGAAAAAATCATGGCATTAGAGAAAGAAATTGCACCAA	3968
QY	1724	TGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTC	1783
Db	3969	TCAACTTTAAAGTTAAACGGTGAATGATTATCCAGTCAGTCTCTCCACTTGACCTGC	4028
QY	1784	AGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGC	1843
Db	4029	ATCCCTCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAATATCGGATGGAACCTTTTAC	4088
QY	1844	AGTGGCCGTGAGGACCGAGTCAGGAGCTGATGAAGCCCAACAGGAGCTTTGGTCCAG	1903
Db	4089	AGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCCAACAGAGATTTTGGACCAT	4148
QY	1904	CATCTCAGCACTTTCTTCCAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAA	1963
Db	4149	CCTCTCAGCACTTTCTCTACGTCACTCCAGCTGCCGTGGCAAGATCCATTTTCACATA	4208
QY	1964	ACAAAGTGCCCTACTATATCAACCAACGAGACTCAAAC	2000
Db	4209	ATAAGTGCCCTATTACATCAACCATCAACACACAGAC	4245

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320

2Y 1404 AGAGAAACTCTACCAAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGATGTAC 1463
Db 844 AGAGGAGCCTCATTTGTAGAGCAAGATACCTCCCGAAACAGCGGATCCAGATCTCAG 903
2Y 1464 TCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCA 1523
Db 904 CCGCTTTGTATGGAAGCAGGCGAGCGGTGGCCAGTGAACTGTGGGAGAAAGTTGACAGCCG 963
2Y 1524 CTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGA 1583
Db 964 CTGTGTGGACCAAGCACCCTCACATTGACGGGACTCTGGAGCAGCTCTTGGAGATTCA--G 1021
2Y 1584 GGCCAQGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCA 1643
Db 1022 GGGCATGGAGGAACCTAAGCACTACTCTGAGCCAAGCTGAGGAGTCCGAGCCACTTGGGA 1081
QY 1644 GCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACT 1703
Db 1082 GCCCATTTGGGATCTCTTCATTGATTCACCTCCAGAGCACATCCAGGCTATTAAAGCTGTT 1141
QY 1704 TCGAGGAGAAATTGGCCCTCTGAAAGAGAAAGTGAGCCACGTCAATGACCTTGTCTGCCA 1763
Db 1142 CAAAGAAGAAATTCCTCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCACCA 1201
QY 1764 GCTTACCACCTTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAA 1823
Db 1202 ACTTGCCATTTCTGATGTGCACCTTGTCAATGGAGAAATCCAGGCCCTTGGAAACAGATCAA 1261
QY 1824 CACCAGATGGAAGCTTCTCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGC 1883
Db 1262 CGTCCGATGGAACAACTACAGGCGTCAGTTCAGTTCAGAGGCTTAAGCAGCTCCAGGATGC 1321
QY 1884 CCACAGGGACTTTGTGTCAGCATCTCAGCACTTTCTTCCAGCTGTGTCCAGGTCCTG 1943
Db 1322 CCACCGGACTTTGGGCTGGTTCACAGCACTTTCTCTCCTCTGTCTCCAGGTTCCCTG 1381
QY 1944 GGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAAC 2000
Db 1382 GGAAGAGCAATTTACCCCAATAAAGTTCCCTACTACATCAACACCAGGCTCAGAC 1438

RESULT 8

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match 4.0%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 8.8e-15;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 465 CCTAAACGCCCAAGTACAAACAATAGGTGCTTCAAGAAGATCTAGAAACAAGAACAGT 524
Db 16 CCTGCAAAACCTGCTTGAAGAACATATAAAGTTTGCAAAAGTGAAGCTCGAAGCTGAGCAGT 75
QY 525 CAGGTCATATTCTCTCACTCACATGTTGGTGGTAGTTGATGAATCTAGTGGAGATCACGC 584
Db 76 GAAGGTGAATTCCTTAACCTCATATGGTGGTGAATTTGGATGAAAAACAGTGGGGAGAGCG 135
QY 585 AACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG 644
Db 136 CACAGCTGTTTGGAGAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 645 ATGGA 649
Db 196 CTGGA 200

RESULT 9

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match 3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGAATCTGAAGACCTAAACGCCCAAGTACAAACAATAGGTGCTTCAAGAAGATCTAGA 512
Db 4 TGACCTGCCCTCCCTGCAGAGCTGCTTCAAGAACATAAAAGTTTGCAAAATGACCTGA 63
QY 513 ACAAGAACAGTCAGGTCAGGTCAATTCTCTCACTCACATGGTGGTGGTAGTGAATCTAG 572
Db 64 AGCTGAACAGGTGAAGGTAAATTCCTTAACCTCACATGGTGGTGAATTTGGATGAAACAG 123
QY 573 TGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGGAGATCGATGGC 632
Db 124 TGGGAGAGTGCCACAGCTCTTCTGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGAC 183
QY 633 AAACATCTGTAGATGGA 649
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 10

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression

FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091.501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 200
TYPE: DNA
ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match 3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 453 TGATCTTGAAGACCTAAAGCCCAAGTACAAACATAAGGTGCTTCAAGAAGATCTAGA 512
Db 4 TGATGTGAATCTCTACAAAGCTGCTAGAAGAACATAAAAGTTTGCAAAGTGATCTGA 63
QY 513 ACAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTGTAGTTGATGAATCTAG 572
Db 64 GGCTGAACAGGTGAAGTAATTCACCTAATCACATGGTGGTCATTGTTGATGAAAACAG 123
QY 573 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 632
Db 124 TGGTGAGAGCGGTACAGCTATCCTAGAGACCAAGTTACAGAAACTTGGTGAGCGCTGGAC 183
QY 633 AAACATCTGTAGATGGA 649
Db 184 AGCAGTATGCCGTTGGA 200

RESULT 11
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match 3.8%; Score 76.6; DB 1; Length 7218;
Best Local Similarity 6.7%; Pred. No. 8.8e-13;
Matches 28; Conservative 236; Mismatches 155; Indels 0; Gaps 0;
QY 113 TTCTAATGATGTGGAAGTGGTGAAGACCAAGTTTCATACTCATGAGGGGTACATGATGG 172
Db 1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRR 1415
QY 173 ATTTGACAGCCCATCAGGCGCGGTTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTG 232
Db 1414 RR 1355
QY 233 GAACAGGAAATTTATCAGAAAGATGAAGAACTGAAGTACAAGACGAGATGAATCTCTAA 292
Db 1354 RR 1295
QY 293 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAG 352
Db 1294 RR 1235
QY 353 TTTTAATGGATCTCCAGAATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG 412
Db 1234 RR 1175
QY 413 AAGAAAGAACAAAGGAAATGGAGGAGAGCTCTTGGACCTGATCTTGAAGACCTAAAC 472
Db 1174 RR 1115
QY 473 GCCAAGTACAACAATAAGGTGCTTCAAGAAGATCTAGAACAAACAAGTCAAGGTC 531
Db 1114 RR 1056

RESULT 12
US-09-687-875A-13
Sequence 13, Application US/09687875A
Patent No. 654786
GENERAL INFORMATION:
APPLICANT: Xiao, Xiao
APPLICANT: Liu, Paul
TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED P
FILE REFERENCE: 00792
CURRENT APPLICATION NUMBER: US/09/687,875A
CURRENT FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/158,868
PRIOR FILING DATE: 1999-10-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.1
SEQ ID NO 13
LENGTH: 238
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match 3.2%; Score 63.6; DB 4; Length 238;

STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
SEQUENCE CHARACTERISTICS:
LENGTH: 1179 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1179
SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match 2.2%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.0058;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy	371	ATCAGAAACTGAAAGACTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAAAA	430
Db	602	AGCAAGAATAAAAGATTGATCGATCAGACAAAGAAAATGGAGATACGATCGGAGGAA	661
Qy	431	TGGAGGAAGAGCCTCTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACACATA	490
Db	662	TTGTAGAAGTGCTCGTTGGAGGCGTTCCAGCTGGATTAGGAAGCTACGTACATGGGACA	721
Qy	491	AGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGG	550
Db	722	CGAAGCTAGATGCCAAAATCGCACAAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCCGTAG	781
Qy	551	TGGTGGTAGTTGATGATCTAGTGG	575
Db	782	AATTGGGGTCGGATTCACTTCTGG	806

Search completed: April 5, 2004, 08:24:49
Job time : 100.041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 04:53:30 ; Search time 485.207 Seconds
(without alignments)
15454.047 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaca 2001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:**
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:**
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:**
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:**
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:**
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:**
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:**
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:**
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:**
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:**
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:**
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:**
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:**
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:**
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:**
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:**
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:**
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2001	100.0	3999	10	US-09-845-416-6 Sequence 6, Appli
2	2001	100.0	4966	10	US-09-845-416-28 Sequence 28, Appli
3	2001	100.0	4990	10	US-09-845-416-34 Sequence 34, Appli
4	1709	85.4	3858	10	US-09-845-416-9 Sequence 9, Appli
5	1709	85.4	4825	10	US-09-845-416-29 Sequence 29, Appli
6	1709	85.4	4848	10	US-09-845-416-35 Sequence 35, Appli
7	1709	85.4	5060	10	US-09-845-416-36 Sequence 36, Appli
8	1562.8	83.1	4182	10	US-09-845-416-2 Sequence 2, Appli
9	1662.8	83.1	5149	10	US-09-845-416-27 Sequence 27, Appli
10	1283	64.1	5462	15	US-10-149-736-41 Sequence 41, Appli
11	1182.6	59.1	8689	15	US-10-149-736-42 Sequence 42, Appli
12	1182.6	59.1	11058	10	US-09-845-416-1 Sequence 1, Appli
13	1182.6	59.1	11443	15	US-10-149-736-44 Sequence 44, Appli
14	1182.6	59.1	12057	15	US-10-149-736-47 Sequence 47, Appli
15	1182.6	59.1	13957	9	US-09-782-378A-22 Sequence 22, Appli

16	1182.6	59.1	13957	9	US-09-880-107-2284	Sequence 2284, Ap
17	1182.6	59.1	13957	15	US-10-149-736-1	Sequence 1, Appli
18	1182.6	59.1	14069	12	US-10-342-887-434	Sequence 434, App
19	1182.6	59.1	14082	12	US-10-342-887-981	Sequence 981, App
20	1182.6	59.1	14082	15	US-10-341-434-108	Sequence 108, App
21	1182	59.1	2169	10	US-09-845-416-4	Sequence 4, Appli
22	1182	59.1	3531	10	US-09-845-416-10	Sequence 10, Appli
23	1182	59.1	4498	10	US-09-845-416-30	Sequence 30, Appli
24	1180	59.0	5339	15	US-10-149-736-40	Sequence 40, Appli
25	1013	50.6	3510	10	US-09-845-416-12	Sequence 12, Appli
26	1013	50.6	4476	10	US-09-845-416-31	Sequence 31, Appli
27	1004	50.2	13815	15	US-10-149-736-2	Sequence 2, Appli
28	835	41.7	1821	10	US-09-845-416-13	Sequence 13, Appli
29	727	36.3	5417	15	US-10-149-736-39	Sequence 39, Appli
30	678.4	33.9	4414	10	US-09-845-416-32	Sequence 32, Appli
31	678	33.9	1991	10	US-09-845-416-7	Sequence 7, Appli
32	677	33.8	1657	10	US-09-845-416-14	Sequence 14, Appli
33	666.4	33.3	3446	10	US-09-845-416-15	Sequence 15, Appli
34	448	22.4	1434	10	US-09-845-416-4	Sequence 4, Appli
35	408.2	20.4	11096	15	US-10-149-736-3	Sequence 3, Appli
36	397	19.8	10302	9	US-09-782-378A-23	Sequence 23, Appli
37	397	19.8	10302	15	US-10-149-736-3	Sequence 3, Appli
38	387	19.3	387	15	US-10-149-736-32	Sequence 32, Appli
39	385	19.2	16531	14	US-10-101-510-667	Sequence 667, App
40	350	17.5	1340	10	US-09-845-416-11	Sequence 11, Appli
41	348	17.4	348	15	US-10-149-736-31	Sequence 31, Appli
42	330	16.5	333	15	US-10-149-736-9	Sequence 9, Appli
43	327	16.3	327	15	US-10-149-736-8	Sequence 8, Appli
44	324	16.2	324	15	US-10-149-736-33	Sequence 33, Appli
45	322.6	16.1	5106	12	US-10-220-120-157	Sequence 157, App

ALIGNMENTS

RESULT 1
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match		100.0%;	Score 2001;	DB 10;	Length 3999;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2001;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTATCAAAACACAGCTTTAGAAGAA	60		
Db	1000	GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTATCAAAACACAGCTTTAGAAGAA	1059		
QY	61	GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTCTTAAT	120		
Db	1060	GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTCTTAAT	1119		
QY	121	GATGTGGAAGTGGTGAACACACAGTTTCATACACTCATGAGGGGTACATGATGGATTGACA	180		
Db	1120	GATGTGGAAGTGGTGAACACACAGTTTCATACACTCATGAGGGGTACATGATGGATTGACA	1179		
QY	181	GCCCATCAGGGCCGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA	240		

Db 1180 GCCCATCAGGCGGGTGGTAATATTCTCAATTGGGAAGTAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAGA 1299
QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 420
Db 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 1419
QY 421 ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 480
Db 1420 ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 1479
QY 481 CAACAAATAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTC 540
Db 1480 CAACAAATAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTC 1539
QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACTTAAGGTATTGGAGATCGATGGGCAAAACATCTGTAGTGGACAGAACCCGC 660
Db 1600 GAACAACTTAAGGTATTGGAGATCGATGGGCAAAACATCTGTAGTGGACAGAACCCGC 1659
QY 661 TGGGTTCTTTACAAGACCAAGCCTGACCTAGTCTCTGGACTGACCACTATTGGAGCCTCT 720
Db 1660 TGGGTTCTTTACAAGACCAAGCCTGACCTAGTCTCTGGACTGACCACTATTGGAGCCTCT 1719
QY 721 CCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db 1720 CCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 1779
QY 781 AAACATAGAAATGCCATCTTCTCTGATGTTGGAGTACCTACTCATAGATTACTGCAACAG 840
Db 1780 AAACATAGAAATGCCATCTTCTCTGATGTTGGAGTACCTACTCATAGATTACTGCAACAG 1839
QY 841 TTCCCTCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAAT 900
Db 1840 TTCCCTCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAAT 1899
QY 901 GTCCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGACTCCAAAGGAGTAAAGAG 960
Db 1900 GTCCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGACTCCAAAGGAGTAAAGAG 1959
QY 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 1020
Db 1960 CTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 2019
QY 1021 AACCTGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
Db 2020 AACCTGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2079
QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACITCGGAAAAAGTCT 1140
Db 2080 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACITCGGAAAAAGTCT 2139
QY 1141 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCT 1200
Db 2140 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCT 2199
QY 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2200 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2259
QY 1261 ATTGAGGCGACTTTCCAGAGATTCCAGAACGACAGACGATGTACATAGGCGCTTCAAGAGG 1320

Db 2260 ATTGAGGCGGACTTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGG 2319
QY 1321 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 2320 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 2379
QY 1381 ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGAGCTGCCTCCT 1440
Db 2380 ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGAGCTGCCTCCT 2439
QY 1441 GAGGAGAGAGCCCGAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 2440 GAGGAGAGAGCCCGAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGGTCAATACT 2499
QY 1501 GAGTGGGAAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 2500 GAGTGGGAAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 2559
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTCAAGCTGGCGCCAAGCT 1620
Db 2560 GAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTCAAGCTGGCGCCAAGCT 2619
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
Db 2620 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCAAGAT 2679
QY 1681 CACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCCTCTGAAAAGAGAACGTGAGC 1740
Db 2680 CACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCCTCTGAAAAGAGAACGTGAGC 2739
QY 1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 1800
Db 2740 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC 2799
QY 1801 CTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGAC 1860
Db 2800 CTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGAC 2859
QY 1861 CGAGTCAGCAGCTGCATGAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 2860 CGAGTCAGCAGCTGCATGAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 2919
QY 1921 TCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCCTACTAT 1980
Db 2920 TCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCCTACTAT 2979
QY 1981 ATCAACCACGAGACTCAAACA 2001
Db 2980 ATCAACCACGAGACTCAAACA 3000

RESULT 2

US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 100.0%; Score 2001; DB 10; Length 4966;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
1	GGCAGTTTCATTTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA	60					
1757	GGCAGTTTCATTTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA	1816					
61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT	120					
1817	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT	1876					
121	GATGTGGAAGTGGTGAAGACCAAGTTCATATCTCATGAGGGGTACATGATGGATTTGACA	180					
1877	GATGTGGAAGTGGTGAAGACCAAGTTCATATCTCATGAGGGGTACATGATGGATTTGACA	1936					
181	GCCCATCAGGGCCGGTTGGTAAATATTCTACAAATTGGGAAGTAAAGCTGATTGGAACAGGA	240					
1937	GCCCATCAGGGCCGGTTGGTAAATATTCTACAAATTGGGAAGTAAAGCTGATTGGAACAGGA	1996					
241	AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA	300					
1997	AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA	2056					
301	TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG	360					
2057	TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG	2116					
361	GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA	420					
2117	GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA	2176					
421	ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	480					
2177	ACAAGGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	2236					
481	CAACAACATAAGGTGCTTCAAGNAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTC	540					
2237	CAACAACATAAGGTGCTTCAAGNAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTC	2296					
541	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA	600					
2297	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA	2356					
601	GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC	660					
2357	GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC	2416					
661	TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGAGTACCACTATTGGAGCCTCT	720					
2417	TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGAGTACCACTATTGGAGCCTCT	2476					
721	CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC	780					
2477	CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC	2536					
781	AAACTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG	840					
2537	AAACTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG	2596					
841	TTCCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT	900					
2597	TTCCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT	2656					
901	GTCCTACAGATGCTACCCGTGAAGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAG	960					
2657	GTCCTACAGATGCTACCCGTGAAGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAG	2716					
961	CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC	1020					
2717	CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC	2776					
1021	AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCCCGATGATCGAGTC	1080					

Db	2777	AACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTC	2836
Qy	1081	CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCT	1140
Db	2837	CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCT	2896
Qy	1141	CTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT	1200
Db	2897	CTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT	2956
Qy	1201	CTGCAGGAACTTCTGTTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT	1260
Db	2957	CTGCAGGAACTTCTGTTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT	3016
Qy	1261	ATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG	1320
Db	3017	ATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG	3076
Qy	1321	GAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTG	1380
Db	3077	GAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTG	3136
Qy	1381	ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGAGCCCCAGAGAGCTGCTCCT	1440
Db	3137	ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGAGCCCCAGAGAGCTGCTCCT	3196
Qy	1441	GAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT	1500
Db	3197	GAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT	3256
Qy	1501	GAGTGGGAAAAATTGAAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT	1560
Db	3257	GAGTGGGAAAAATTGAAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT	3316
Qy	1561	GAAAGACTCCAGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGGGCCAAGCT	1620
Db	3317	GAAAGACTCCAGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGGGCCAAGCT	3376
Qy	1621	GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGAT	1680
Db	3377	GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGAT	3436
Qy	1681	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCTCTGAAAGAGAAACGTGAGC	1740
Db	3437	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCTCTGAAAGAGAAACGTGAGC	3496
Qy	1741	CACGTCAATGACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC	1800
Db	3497	CACGTCAATGACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC	3556
Qy	1801	CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC	1860
Db	3557	CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGAC	3616
Qy	1861	CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT	1920
Db	3617	CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT	3676
Qy	1921	TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTAT	1980
Db	3677	TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTAT	3736
Qy	1981	ATCAACCACGAGACTCAAACA 2001	
Db	3737	ATCAACCACGAGACTCAAACA 3757	

RESULT 3
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-34

Query Match 100.0%; Score 2001; DB 10; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1840
QY 61 GTAATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCATAAT 120
Db 1841 GTAATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCATAAT 1900
QY 121 GATGTGGAAGTGGTGAAGAACCCAGTTTCTATCTATCATGAGGGGTACATGATGGATTTGACA 180
Db 1901 GATGTGGAAGTGGTGAAGAACCCAGTTTCTATCTATCATGAGGGGTACATGATGGATTTGACA 1960
QY 181 GCCCATCAGGGCCGGGTGGTAATATTTCTACAAATGGGAAGTAAGTCGATTGGAAACAGGA 240
Db 1961 GCCCATCAGGGCCGGGTGGTAATATTTCTACAAATGGGAAGTAAGTCGATTGGAAACAGGA 2020
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 300
Db 2021 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 2080
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTTAATG 360
Db 2081 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTTAATG 2140
QY 361 GATCTCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 420
Db 2141 GATCTCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 2200
QY 421 ACAAGGAAATGGAGGAAGAGCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 2201 ACAAGGAAATGGAGGAAGAGCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2260
QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTC 540
Db 2261 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTC 2320
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 2321 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
Db 2381 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 2440
QY 661 TGGGTTCTTTTACAAGACCGCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 720
Db 2441 TGGGTTCTTTTACAAGACCGCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 2500
QY 721 CCTACTCAGACTGTTACTCTGTGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db 2501 CCTACTCAGACTGTTACTCTGTGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 2560
QY 781 AAACCTAGAAATGCCATCTTCCITGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840

Db 2561 AAACCTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATATAGATTACTGCAACAG 2620
QY 841 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTTGTGCTTACAGAAAGCTGAAACAACTGCCAAT 900
Db 2621 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTTGTGCTTACAGAAAGCTGAAACAACTGCCAAT 2680
QY 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAG 960
Db 2681 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAGAG 2740
QY 961 CTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 1020
Db 2741 CTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 2800
QY 1021 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTC 1080
Db 2801 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTC 2860
QY 1081 CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCT 1140
Db 2861 CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCT 2920
QY 1141 CTCACATATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCTGTCACCTTTCT 1200
Db 2921 CTCACATATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCTGTCACCTTTCT 2980
QY 1201 CTGAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACCT 1260
Db 2981 CTGAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACCT 3040
QY 1261 ATTGGAGGCGACTTTCAGCAGATTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
Db 3041 ATTGGAGGCGACTTTCAGCAGATTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3100
QY 1321 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 3101 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3160
QY 1381 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGCTGCCTCCT 1440
Db 3161 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGCTGCCTCCT 3220
QY 1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1500
Db 3221 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 3280
QY 1501 GAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 3281 GAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3340
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGGCCCAAGCT 1620
Db 3341 GAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGGCCCAAGCT 3400
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGAT 1680
Db 3401 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGAT 3460
QY 1681 CACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCCTCTGAAAAGAGACGTGAGC 1740
Db 3461 CACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGGCCCTCTGAAAAGAGACGTGAGC 3520
QY 1741 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
Db 3521 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 3580
QY 1801 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
Db 3581 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 3640
QY 1861 CGAGTCAGGACGTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 3641 CGAGTCAGGACGTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTT 3700

QY	1681	CACCTCGAGAAAGTCAAGGCAC	TTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGC	1741
Db	2539	CACCTCGAGAAAGTCAAGGCAC	TTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGC	2598
QY	1741	CACGTCGAATGACCTTGCTCGCCAGCTT	ACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC	1800
Db	2599	CACGTCGAATGACCTTGCTCGCCAGCTT	ACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC	2658
QY	1801	CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTT	CTGCAGGTGGCCGTCGAGGAC	1860
Db	2659	CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTT	CTGCAGGTGGCCGTCGAGGAC	2718
QY	1861	CGAGTCAGGCAGCTGCATGAAGCCCA	CAGGGACCTTGGTCCAGCATCTCAGCACCTTTCTT	1920
Db	2719	CGAGTCAGGCAGCTGCATGAAGCCCA	CAGGGACCTTGGTCCAGCATCTCAGCACCTTTCTT	2778
QY	1921	TCCACGTCGTCCAGGTCCTTGGGAGAGAGCCATCT	TCGCCAAACAAAGTGCCTACTAT	1980
Db	2779	TCCACGTCGTCCAGGTCCTTGGGAGAGAGCCATCT	TCGCCAAACAAAGTGCCTACTAT	2838
QY	1981	ATCAACCAACGAGACTCAAACA	2001	
Db	2839	ATCAACCAACGAGACTCAAACA	2859	

RESULT 5

RESOL 3
US-09-845-416-29

03-03-043-110-23
: Sequence 29. Application US/09845416

; sequence 23, Application 03/038
; Publication No. US20030171312A1; PUBLICATION NO: US20
: GENERAL INFORMATION:; GENERAL INFORMATION:
; APPLICANT: XTAO. XTAO

APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MTINGENE AND USE

1. TITLE OF INVENTION: DNA SEQUENCING

1. TITLE OF INVENTION: THEREOF
2. FILE REFERENCE: DE1142

FILE REFERENCE: DELI42
CURRENT APPLICATION NUMBER: IIS/O

;; CURRENT APPLICATION NUMBER: US/09/84
: CURRENT FILING DATE: 2001-04-30

; CURRENT FILING DATE: 2001-04
 : PRIOR APPLICATION NUMBER: 60/

; PRIOR APPLICATION NUMBER
 : PRIOR FILING DATE: 2000-

;; PRIOR FILING DATE: 2000-04-
NUMBER OF SEQ ID NOS: 36

; NUMBER OF SEQ
: SOFTWARE: DEF

```

; SOFTWARE: Pat
; GEO ID NO 38

```

```

; SEQ ID NO 2
      LENGTH: 4

```

```

; LENGTH: 4825
; TYPE: DNA

```

TYPE: DNA

Query Match	85.4%	Score 1709;	DB 10;	Length 4825;
Best Local Similarity	93.0%;	Pred. No. 0;		
Matches 1860;	Conservative	0;	Mismatches	0;
			Indels	141;
			Gaps	1;

Db	2057	TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG	2111
Qy	361	GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAAGA	420
Db	2117	GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAAGA	2176
Qy	421	ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTA	480
Db	2177	ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTA	2236
Qy	481	CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC	540
Db	2237	CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC	2296
Qy	541	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA	600
Db	2297	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA	2356
Qy	601	GAACAACTTAAGGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCG	660
Db	2357	GAACAACTTAAGGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCG	2416
Qy	661	TGGGTTCTTTTACAAGACACGACCTGACCTAGTCTCTGGACTGACCACATATTGGAGCCTCT	720
Db	2417	TGGGTTCTTTTACAAGAC	2434
Qy	721	CCTACTCAGACTGTTACTCTCTGGTGACACAACCTGTGGTTACTAAGGAACTGCCATCTCC	780
Db	2435		2434
Qy	781	AAACTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG	840
Db	2435		2455
Qy	841	TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACATGCGCAAT	900
Db	2456	TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACATGCGCAAT	2515
Qy	901	GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAAAGAG	960
Db	2516	GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAAAGAG	2575
Qy	961	CTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC	1020
Db	2576	CTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC	2635
Qy	1021	AACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC	1080
Db	2636	AACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC	2695
Qy	1081	CTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAGAAAGTCT	1140
Db	2696	CTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAGAAAGTCT	2755
Qy	1141	CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT	1200
Db	2756	CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT	2815
Qy	1201	CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT	1260
Db	2816	CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT	2875
Qy	1261	ATTGGAGGCGACTTCCAGCAGTTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGG	1320
Db	2876	ATTGGAGGCGACTTCCAGCAGTTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGG	2935
Qy	1321	GAATTGAAACCTAAAGAACCTGTAACTCATGAGTACTCTTTGAGACTGTACGAATATTTCTG	1380
Db	2936	GAATTGAAACCTAAAGAACCTGTAACTCATGAGTACTCTTTGAGACTGTACGAATATTTCTG	2995
Qy	1381	ACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCTCCT	1440
Db	2996	ACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCTCCT	3055

1200 1020 2870 1080 2930 1140 2990 1200 3050 1260 3110 1320 3170 1380 3230 1440 3290 1500 3350 1560 3410 1620 3470 1680 3530 1740 3590 1800 3650 1860 3710 1920 3770 1980 3830 2001 3851

CTGATGAAACAATGGCAAGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC
CTGATGAAACAATGGCAAGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC
AACCTGGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTC
AACCTGGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTC
CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT
CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT
CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTTGACCAGTGGAAAGCTTCGACCTTTCT
CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCT
CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCT
ATTGGAGCGCACTTCCAGCACTTCCAGCAAGCAGACGATGTACATAGGCGCTTCAAGAGG
ATTGGAGCGCACTTCCAGCACTTCCAGCAAGCAGACGATGTACATAGGCGCTTCAAGAGG
GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG
GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG
ACAGAGCAGCCTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCT
ACAGAGCAGCCTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCT
GAGGAGAGAGCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCT
GAGGAGAGAGCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCT
GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT
GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT
CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGC
CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGC
CAGCTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC
CAGCTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC
CTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGAC
CTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGAC
CGAGTCAAGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTTCTT
CGAGTCAAGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTTCTT
TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT
TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT
ATCAACACAGGAGACTCAAACA 2001
ATCAACACAGGAGACTCAAACA 3851

RESULT 8
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 83.1%; Score 1662.8; DB 10; Length 4182;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTCATCTACTGAGGGTACATGATGGATTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCAAGTTCATCTCATGAGGGTACATGATGGATTGACA 1179
QY 181 GCCCATCAGGCGCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1180 GCCCATCAGGCGCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGA 1299
QY 301 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 420
Db 1360 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 1419
QY 421 ACAAGGAAAAATGGAGGAAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 480
Db 1420 ACAAGGAAAAATGGAGGAAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAGTA 1479
QY 481 CAACAACATAAGTGTCTTCAAGAAGATCTAGAAACAGAAACAAAGTCAAGGTCAATTCTCTC 540
Db 1480 CAACAACATAAGTGTCTTCAAGAAGATCTAGAAACAGAAACAAAGTCAAGGTCAATTCTCTC 1539
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGC 660
Db 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGC 1659
QY 661 TGGGTTCTTTTACAAGACCAAGCCCTGAC-----CTAGCTCCTGGACTGACCACT 708
Db 1660 TGGGTTCTTTTACAAGACATCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 1719

QY	709	ATTGAGCCTCTCTACTCAGACTGTTACTCTGGTGACACA	-----ACCTGTGGTT	759
Db	1720	TTTAGTGCAATGGCTTTTCAGAAAAAAGAAGATGCAGTGAA	CAAGATTTCACACAACCTGGCTTT	1779
QY	760	ACTAAGGAAACTGCCATCTC	-----	779
Db	1780	AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAAC	TGGCCGTTTTTAAACCGGATCTA	1839
QY	780	-----	-----CAAACTAGAAATGCCATCTTCC	801
Db	1840	GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACT	CAAAACAAGATCTTCTTTCAACA	1899
QY	802	TTGATGTTGGAG	-----	813
Db	1900	CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGG	CTGGATAACTTTGCCCGGTGT	1959
QY	814	-----	-----GTACCTACTCATAGATTACTGCAA	837
Db	1960	TGGGATAATTTAGTCCAAAAA	CTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA	2019
QY	838	CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCT	TACAGAAAGCTGAAACAACACTGCC	897
Db	2020	CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCT	TACAGAAAGCTGAAACAACACTGCC	2079
QY	898	AATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGA	AGACTCCCAAGGAGTAAAA	957
Db	2080	AATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGA	AGACTCCCAAGGAGTAAAA	2139
QY	958	GAGCTGATGAAACAATGGCAAGACCTCCAAGTGAAATTA	AGCTCACACAGATGTTTAT	1017
Db	2140	GAGCTGATGAAACAATGGCAAGACCTCCAAGTGAAATTA	AGCTCACACAGATGTTTAT	2199
QY	1018	CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCC	CTGGAAGTTCCGATGATGCA	1077
Db	2200	CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCC	CTGGAAGTTCCGATGATGCA	2259
QY	1078	GTCCTGTTACAAAGACGTTTGGATAACATGAACITTC	AAAGTGAGTGAACTTCGAAAAAG	1137
Db	2260	GTCCTGTTACAAAGACGTTTGGATAACATGAACITTC	AAAGTGAGTGAACTTCGAAAAAG	2319
QY	1138	TCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACC	AGTGAAGCGTCTGCACCTT	1197
Db	2320	TCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACC	AGTGAAGCGTCTGCACCTT	2379
QY	1198	TCTCTGCAGGAACITCTGGTGTGGCTACAGCTGAAAG	ATGATGAATTAAGCCGCGCAGGCA	1257
Db	2380	TCTCTGCAGGAACITCTGGTGTGGCTACAGCTGAAAG	ATGATGAATTAAGCCGCGCAGGCA	2439
QY	1258	CCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAG	AACGATGTACATAGGCGCTTCAAG	1317
Db	2440	CCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAG	AACGATGTACATAGGCGCTTCAAG	2499
QY	1318	AGGGAATTGAAAACATAAAGAACCTGTAAATCATG	AGTACTCTTGAGACTGTACGAATATTT	1377
Db	2500	AGGGAATTGAAAACATAAAGAACCTGTAAATCATG	AGTACTCTTGAGACTGTACGAATATTT	2559
QY	1378	CTGACAGAGCAGCTTTTGGAAAGGACTAGAGAA	ACTCTACAGGAGCCAGAGAGCTGCCT	1437
Db	2560	CTGACAGAGCAGCTTTTGGAAAGGACTAGAGAA	ACTCTACAGGAGCCAGAGAGCTGCCT	2619
QY	1438	CCTGAGGAGAGAGCCACAGAAATGTCACTCGGCTT	CTACGAAAGCAGGCTGAGGAGGTCAAT	1497
Db	2620	CCTGAGGAGAGAGCCACAGAAATGTCACTCGGCTT	CTACGAAAGCAGGCTGAGGAGGTCAAT	2679
QY	1498	ACTGAGTGGAAAAAATTGAACTCGCTGACTCCGCT	GATGGCAGAGAAAAATAGATGAGACC	1557
Db	2680	ACTGAGTGGAAAAAATTGAACTCGCTGACTCCGCT	GATGGCAGAGAAAAATAGATGAGACC	2739
QY	1558	CTTGAAAGACTCCAGGAACCTCAAGAGGCCACGG	ATGAGCTTGACCTCAAGCTGCGCCAA	1617
Db	2740	CTTGAAAGACTCCAGGAACCTCAAGAGGCCACGG	ATGAGCTTGACCTCAAGCTGCGCCAA	2799

RESIST, 9

US-09-845-416-27

: Sequence 27. Application US/09845416

Sequence 2, Application 03/030
: Publication No: US20030171312A1

GENERAL INFORMATION:

APPLICANT: XIAO. XIAO

TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

INVENTOR:	TITLE OF INVENTION:	DATE OF INVENTION:	DATE OF INVENTION:
JOHN D. BROWN	NEW METHOD OF	1900	1900

FILE OF INVENTION: I
FILE REFERENCE: DE1142

```

; FILE REFERENCE: DE1142
: CURRENT APPLICATION NUMBER: US/09/845.416

```

; CURRENT APPLICATION NUMBER: US/0
 : CURRENT FILING DATE: 2001-04-30

; CURRENT FILING DATE: 2001-04-30
 : PRIOR APPLICATION NUMBER: 60/200,777

! PRIOR APPLICATION NUMBER: 607
: PRIOR FILING DATE: 2000-04-28

; PRIOR FILING DATE: 2000-
 : NUMBER OF SEQ ID NOS: 36

```

; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver 2.1

```

SOFTWARE: E
ID NO: 27

```

; SEQ ID NO 2 /
: LENGTH: 5149

```

```

; LENGTH: 5
; TYPE: DNA

```

TYPE: DNA
ORGANISM: Homo sapiens

ORGANISM: HO
118-08-84E-436-27

Query Match 83.1%; Score 1662.8; DB 10; Length 5149;

```

Query Match: 00.19,
Best Local Similarity 88.3%,
Pred, No. 0: 00.19,
Pred, No. 0: 00.19,

```

Best local similarity 80.35; 120: no. of
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY	1	GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA	60
Db	1757	GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA	1816
QY	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTTAAT	120
Db	1817	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTTAAT	1876
QY	121	GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACA	180
Db	1877	GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACA	1936
QY	181	GCCCATCAGGGCCGGTTGGTAAATATTCACATTCGGAAGTAAGCTGATTGGAAACAGGA	240
Db	1937	GCCCATCAGGGCCGGTTGGTAAATATTCACATTCGGAAGTAAGCTGATTGGAAACAGGA	1996
QY	241	AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA	300
Db	1997	AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA	2056

2Y 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
2b |||||
2Y 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2116
2b |||||
2Y 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGCTAACAACAAAACAGAAGAAAGA 420
2b |||||
2Y 2117 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGCTAACAACAAAACAGAAGAAAGA 2176
2b |||||
2Y 421 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
2b |||||
2Y 2177 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2236
2b |||||
2Y 481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540
2b |||||
2Y 2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 2296
2b |||||
2Y 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
2b |||||
2Y 2297 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356
2b |||||
2Y 601 GAACAACCTAAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGC 660
2b |||||
2Y 2357 GAACAACCTAAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGC 2416
2b |||||
2Y 661 TGGGTTCTTTTACAAGACACCAGCCTGAC-----CTAGCTCCTGGACTGACCACT 708
2b |||||
2Y 2417 TGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 2476
2b |||||
2Y 709 ATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACA-----ACCTGTGGTT 759
2b |||||
2Y 2477 TTTAGTGATGGCTTTTCAGAAAAGAAAGATGCAGTGGCAACACATCTGTAGATGGACAGAAGACCGC 2536
2b |||||
2Y 760 ACTAAGGAAAACCTGCCATCTC-----779
2b |||||
2Y 2537 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAACTGGCCGCTTTTAAAGCGGATCTA 2596
2b |||||
2Y 780 -----CAAAGTAAAGTCCATCTTCC 801
2b |||||
2Y 2597 GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 2656
2b |||||
2Y 802 TTGATGTTGGAG-----813
2b |||||
2Y 2657 CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 2716
2b |||||
2Y 814 -----GTACCTACTCATAGATTACTGCAA 837
2b |||||
2Y 2717 TGGGATAAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
2b |||||
2Y 838 CAGTTCCCTCCCTGGACCTGGAAGAGTTTCTTGCTGCTGCTTACAGAAGCTGAAACAACTGCC 897
2b |||||
2Y 2777 CAGTTCCCTCCCTGGACCTGGAAGAGTTTCTTGCTGCTGCTTACAGAAGCTGAAACAACTGCC 2836
2b |||||
2Y 898 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAA 957
2b |||||
2Y 2837 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAA 2896
2b |||||
2Y 958 GAGCTGATGAAACAATGSCAAGACCTCCAAGGTGAAATTTGAAGCTCACAGATGTTTAT 1017
2b |||||
2Y 2897 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACAGATGTTTAT 2956
2b |||||
2Y 1018 CACAACCTGGATGAAAACAGSCAAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 1077
2b |||||
2Y 2957 CACAACCTGGATGAAAACAGSCAAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 3016
2b |||||
2Y 1078 GTCCTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAG 1137
2b |||||
2Y 3017 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAG 3076
2b |||||
2Y 1138 TCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 1197
2b |||||
2Y 3077 TCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 3136
2b |||||

QY 1198 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1257
Db |||||
QY 3137 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 3196
Db |||||
QY 1258 CCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAG 1317
Db |||||
QY 3197 CCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAG 3256
Db |||||
QY 1318 AGGGAATTGAAAACCTTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAAATATT 1377
Db |||||
QY 3257 AGGGAATTGAAAACCTTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAAATATT 3316
Db |||||
QY 1378 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACTCTACCAGAGCCCAGAGAGCTGCCT 1437
Db |||||
QY 3317 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACTCTACCAGAGCCCAGAGAGCTGCCT 3376
Db |||||
QY 1438 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAAT 1497
Db |||||
QY 3377 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAAT 3436
Db |||||
QY 1498 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db |||||
QY 3437 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
Db |||||
QY 1558 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 1617
Db |||||
QY 3497 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 3556
Db |||||
QY 1618 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAA 1677
Db |||||
QY 3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAA 3616
Db |||||
QY 1678 GATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTTGCGCCTCTGAAAAGAGAACGTG 1737
Db |||||
QY 3617 GATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTTGCGCCTCTGAAAAGAGAACGTG 3676
Db |||||
QY 1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 1797
Db |||||
QY 3677 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 3736
Db |||||
QY 1798 AACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1857
Db |||||
QY 3737 AACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 3796
Db |||||
QY 1858 GACCGAGTCAGGCACTGTCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTT 1917
Db |||||
QY 3797 GACCGAGTCAGGCACTGTCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTT 3856
Db |||||
QY 1918 CTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTAC 1977
Db |||||
QY 3857 CTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTAC 3916
Db |||||
QY 1978 TATATCAACCACGAGACTCAAAACA 2001
Db |||||
QY 3917 TATATCAACCACGAGACTCAAAACA 3940
Db |||||

RESULT 10
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; PRIOR FILING DATE: 2002-06-17
; CURRENT APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

```

	Query Match	64.1%;	Score 1283;	DB 15;	Length 5462;				
	Best Local Similarity	99.3%;	Pred. No. 0;						
	Matches 1302;	Conservative	0;	Mismatches	0;	Indels	9;	Gaps	1;
Qy	691	GCTCCTGGACTGACCACTATTGGAGCCCTCTCCCTACTCAGACTGTTTACTCTCGTGACACAA	750						
Dd	1547	GCTCCTGGACTGACCACTATTGGAGCCCTCTCCCTACTCAGACTGTTTACTCTCGTGACACAA	1606						
Qy	751	CCTGTGGTTACTAAGGAAACTGCCATCTCCAACCTAGAAAATGCCATCTTCCCTTGATGTTG	810						
Dd	1607	CCTGTGGTTACTAAGGAAACTGCCATCTCCAACCTAGAAAATGCCATCTTCCCTTGATGTTG	1666						
Qy	811	GAGGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCCTGGAAAAGTTTCTTGCC	870						
Dd	1667	GAG-----CATAGATTACTGCAACAGTTCCCCCTGGACCCTGGAAAAGTTTCTTGCC	1717						
Qy	871	TGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTCTACAGGATGCTACCCCCTAAGGAAGG	930						
Dd	1718	TGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTCTACAGGATGCTACCCCCTAAGGAAGG	1777						

Qy	931	CTCCTAGAA	GACTCCAAGGGAGTAAAAGAGCTGATGAACAACATGGCAAGACCTCCAAAGGT	990
Db	1778	CTCCTAGAAG	ACTCCAAGGGAGTAAAAGAGCTGATGAACAACATGGCAAGACCTCCAAAGGT	1837
Qy	991	GAAATTGAAG	CTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTG	1050
Db	1838	GAAATTGAAG	CTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTG	1897
Qy	1051	AGATCCCCTG	GGAAGGTTCCGATGATGCAGTCCTGTACAAAGACGTTTTGGATAACATGAAC	1110
Db	1898	AGATCCCCTG	GGAAGGTTCCGATGATGCAGTCCTGTACAAAGACGTTTTGGATAACATGAAC	1957
Qy	1111	TTCAAGTGGA	GTGAACCTTCGGAAAAAGTCTCTCAAACATTAGGTCCCCATTGTGAAGCCAGT	1170
Db	1958	TTCAAGTGGA	GTGAACCTTCGGAAAAAGTCTCTCAAACATTAGGTCCCCATTGTGAAGCCAGT	2017
Qy	1171	TCTGACCAG	TGSAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGCTACAGCTG	1230
Db	2018	TCTGACCAG	TGSAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGCTACAGCTG	2077
Qy	1231	AAAGATGAT	GAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCAGCAGTTTCAGAAG	1290
Db	2078	AAAGATGAT	GAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCAGCAGTTTCAGAAG	2137
Qy	1291	CAGAACGAT	GTACATAGGGCCCTTCAAGAGGGGAATTGAAAAACTAAAGAACCTGTAAATCATG	1350
Db	2138	CAGAACGAT	GTACATAGGGCCCTTCAAGAGGGGAATTGAAAAACTAAAGAACCTGTAAATCATG	2197
Qy	1351	AGTACTCTT	GACACTGTACGAATATTTCTGACACAGCAGCCCTTTGGAAGGACTAGAGAAA	1410
Db	2198	AGTACTCTT	GACACTGTACGAATATTTCTGACACAGCAGCCCTTTGGAAGGACTAGAGAAA	2257
Qy	1411	CTCTACCAG	GAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTT	1470
Db	2258	CTCTACCAG	GAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTT	2317
Qy	1471	CTACGAAAG	CAGCGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTTGCACTCCGCT	1530
Db	2318	CTACGAAAG	CAGCGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTTGCACTCCGCT	2377
Qy	1531	GACTGGCAG	AGAAAAATAGATGAGACCCCTTGAAAAAGACTCCAGGAACCTTCAAGAGGCCACG	1590
Db	2378	GACTGGCAG	AGAAAAATAGATGAGACCCCTTGAAAAAGACTCCAGGAACCTTCAAGAGGCCACG	2437

RESULT 11
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

	Query Match	59.1%;	Score 1182.6;	DB 15;	Length 8689;
	Best Local Similarity	99.7%;	Pred. No. 0;		
	Matches 1185;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	813	GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCCCTG	872		
DB	2992	GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCCCTG	3051		
QY	873	GTTTACAGAAGCTGAAACAACCTGCCAATGTCTCTACAGGATGCTATCCCGTAAGGAAAGGCT	932		
DB	3052	GTTTACAGAAGCTGAAACAACCTGCCAATGTCTCTACAGGATGCTATCCCGTAAGGAAAGGCT	3111		
QY	933	CCTAGAAGACTCCAAAGGGAGTAAAGAGCTGTGATGAAACAATGGCAAGACCTCCCAAGGTGA	992		
DB	3112	CCTAGAAGACTCCAAAGGGAGTAAAGAGCTGTGATGAAACAATGGCAAGACCTCCCAAGGTGA	3171		
QY	993	AATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAAAACAGCCCAAAAAATCCTGAG	1052		
DB	3172	AATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAAAACAGCCCAAAAAATCCTGAG	3231		

```
1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
Db |||||
3232 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 3291
2Y |||||
1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCCAGTTC 1172
Db |||||
3292 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCCAGTTC 3351
2Y |||||
1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db |||||
3352 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3411
2Y |||||
1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 1292
Db |||||
3412 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 3471
2Y |||||
1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
Db |||||
3472 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 3531
2Y |||||
1353 TACTCTTGACACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
Db |||||
3532 TACTCTTGACACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 3591
2Y |||||
1413 CTACCAAGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db |||||
3592 CTACCAAGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 3651
2Y |||||
1473 ACGAAAGCAGGCTGAGGAGGTCAATACAGTGGGAAAAATTTGAACCTGCACTCCGCTGA 1532
Db |||||
3652 ACGAAAGCAGGCTGAGGAGGTCAATACAGTGGGAAAAATTTGAACCTGCACTCCGCTGA 3711
2Y |||||
1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
Db |||||
3712 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 3771
2Y |||||
1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 1652
Db |||||
3772 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 3831
2Y |||||
1653 CGATCTCCTCAATTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db |||||
3832 CGATCTCCTCAATTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3891
2Y |||||
1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATATGACCTTGCTCGCCAGCTTACCAC 1772
Db |||||
3892 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATATGACCTTGCTCGCCAGCTTACCAC 3951
2Y |||||
1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATG 1832
Db |||||
3952 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATG 4011
2Y |||||
1833 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 1892
Db |||||
4012 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 4071
2Y |||||
1893 CTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
Db |||||
4072 CTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 4131
2Y |||||
1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
Db |||||
4132 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4180
```

RESULT 12
US-09-845-416-1
; Sequence 1, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-1

Query Match	59.1%;	Score	1182.6;	DB	10;	Length	11058;
Best Local Similarity	99.7%;	Pred. No.	0;				
Matches	1185;	Conservative	0;	Mismatches	4;	Indels	0;
Gaps	0;						

QY	813	GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTTGAAAAAGTTTCTTGCCTG	872
Db	8052	GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTTGAAAAAGTTTCTTGCCTG	8111
QY	873	GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGTCACCCGTAAGGAAAGGCT	932
Db	8112	GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGGATGTCACCCGTAAGGAAAGGCT	8171
QY	933	CCTAGAAGACTCCTCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA	992
Db	8172	CCTAGAAGACTCCTCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGA	8231
QY	993	AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCCAAAAAATCCTGAG	1052
Db	8232	AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCCAAAAAATCCTGAG	8291
QY	1053	ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAAGACGTTTGGATAACATGAACCTT	1112
Db	8292	ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAAGACGTTTGGATAACATGAACCTT	8351
QY	1113	CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTC	1172
Db	8352	CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTC	8411
QY	1173	TGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	1232
Db	8412	TGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	8471
QY	1233	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTCCAGCAGTTTCAGAAGCA	1292
Db	8472	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTCCAGCAGTTTCAGAAGCA	8531
QY	1293	GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG	1352
Db	8532	GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG	8591
QY	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACT	1412
Db	8592	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACT	8651
QY	1413	CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT	1472
Db	8652	CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT	8711
QY	1473	ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA	1532
Db	8712	ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA	8771
QY	1533	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACCGGA	1592
Db	8772	CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACCGGA	8831
QY	1593	TGAGCTGGACCTCAAGCTGCGCCAAAGCTAGGTGATCAAGGGATCTCTGGCAGCCCGTGGG	1652
Db	8832	TGAGCTGGACCTCAAGCTGCGCCAAAGCTAGGTGATCAAGGGATCTCTGGCAGCCCGTGGG	8891


```

; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 12057
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
JS-10-149-736-47

Query Match          59.1%; Score 1182.6; DB 15; Length 12057;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2Y 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGCCTG 8319

2Y 873 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8320 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

2Y 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

2Y 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 8499

2Y 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 8559

2Y 1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTTC 1172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8560 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTTC 8619

2Y 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCGACGAACTTCTGGTGTGGCTACAGCTGAA 1232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8620 TGACCAGTGAAGCGTCTGCACCTTTCTCGACGAACTTCTGGTGTGGCTACAGCTGAA 8679

2Y 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAAGCA 1292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8680 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAAGCA 8739

2Y 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8740 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG 8799

2Y 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 8859

2Y 1413 CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 1472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8860 CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 8919

2Y 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACTCCGCTGA 1532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACTCCGCTGA 8979

2Y 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039

2Y 1593 TGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGG 1652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

Db 9040 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9099
QY 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGA 1712
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGA 9159
QY 1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCCCAGCTTACCAC 1772
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9160 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCCCAGCTTACCAC 9219
QY 1773 TTTGGGCATTGAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9220 TTTGGGCATTGAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 9279
QY 1833 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9280 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 9339
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 1952
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 9399
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 2001
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 9448

RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match          59.1%; Score 1182.6; DB 9; Length 13957;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCTG 872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCTG 8319

QY 873 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8320 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

QY 933 CTAGAAGACTCCAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8380 CTAGAAGACTCCAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8440 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 8499

QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8500 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 8559
```

QY	1113	CAAGTGGAGTGAACTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC	1172
Db	8560	CAAGTGGAGTGAACTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC	8619
QY	1173	TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCGGTGTGGCTACAGCTGAA	1232
Db	8620	TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCGGTGTGGCTACAGCTGAA	8679
QY	1233	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA	1292
Db	8680	AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA	8739
QY	1293	GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTTAAGAACCTTGTAATCATGAG	1352
Db	8740	GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTTAAGAACCTTGTAATCATGAG	8799
QY	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT	1412
Db	8800	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT	8859
QY	1413	CTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT	1472
Db	8860	CTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT	8919
QY	1473	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAACCTTGCACTCCGCTGA	1532
Db	8920	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAACCTTGCACTCCGCTGA	8979
QY	1533	CTGGCAGAGAAATAAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA	1592
Db	8980	CTGGCAGAGAAATAAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA	9039
QY	1593	TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG	1652
Db	9040	TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG	9099
QY	1653	CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA	1712
Db	9100	CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA	9159
QY	1713	AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	1772
Db	9160	AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC	9219
QY	1773	TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG	1832
Db	9220	TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG	9279
QY	1833	GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGA	1892
Db	9280	GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGA	9339
QY	1893	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGC	1952
Db	9340	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGC	9399
QY	1953	CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA	2001
Db	9400	CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA	9448

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:52:43 ; Search time 3498.53 Seconds
(without alignments)
17079.796 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcatgtgagag.....tcaaccacgagactcaaac 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678.4	33.9	3870	11 BC036103	BC036103 Homo sapi
2	678.4	33.9	5691	29 AY399453	AY399453 Homo sapi
3	591.8	29.6	834	12 BI729851	BI729851 603349511
4	547.4	27.4	3056	11 AK044536	AK044536 Mus muscu

5	547.4	27.4	5697	29 AY399455	AY399455 Mus muscu
6	491.4	24.6	750	12 BI730168	BI730168 603349711
7	486.6	24.3	579	9 AL121550	AL121550 DKFZp762L
8	471.6	23.6	728	14 CB228986	CB228986 AGENCOURT
9	459.4	23.0	1047	14 CB850319	CB850319 MRA-0070
10	446	22.3	5676	29 AY399454	AY399454 Pan trogl
11	424.2	21.2	663	12 BM488464	BM488464 pgm2n.pk0
12	409	20.4	595	14 CB177816	CB177816 is21c01.x
13	389	19.4	644	13 BU313510	BU313510 603540290
14	324.6	16.2	402	14 CB547284	CB547284 AMGNNUC:S
15	310.8	15.5	2874	29 AY408546	AY408546 Homo sapi
16	309.8	15.5	3753	11 AK081426	AK081426 Mus muscu
17	301.2	15.1	2874	29 AY408548	AY408548 Mus muscu
18	278.2	13.9	9915	29 AY407022	AY407022 Homo sapi
19	265.4	13.3	1122	14 CF109978	CF109978 Shultzomi
20	265.4	13.3	3051	11 BC036095	BC036095 Homo sapi
21	261.2	13.1	2334	11 BC011062	BC011062 Mus muscu
22	261.2	13.1	9691	29 AY407024	AY407024 Mus muscu
23	257.8	12.9	1541	11 AK034383	AK034383 Mus muscu
24	255.4	12.8	9096	29 AY407023	AY407023 Pan trogl
25	252.4	12.6	772	12 BI250598	BI250598 602993659
26	233.6	11.7	826	14 CD656896	CD656896 AGENCOURT
27	233	11.6	835	12 BI553820	BI553820 603190772
28	230.8	11.5	784	12 EG212445	EG212445 RST32032
29	230.2	11.5	763	14 CB518960	CB518960 UI-M-GH0-
30	229.2	11.5	778	14 CB524596	CB524596 UI-M-FY0-
31	226	11.3	250	10 BF963618	BF963618 QV2-NN004
32	222.2	11.1	2874	29 AY408547	AY408547 Pan trogl
33	216.6	10.8	681	13 BU301653	BU301653 603609005
34	199.8	10.0	696	12 BU075057	BU075057 BJ075057
35	199.2	10.0	697	13 BU456556	BU456556 603771141
36	184.4	9.2	812	14 CF748008	CF748008 UI-M-HE0-
37	182.4	9.1	645	10 AW467977	AW467977 he29g08.x
38	179.6	9.0	466	14 CD549993	CD549993 B0305E01-
39	172.2	8.6	630	9 AL855376	AL855376 AL855376
40	167.6	8.4	502	9 AL602076	AL602076 DKFZp313B
41	167.2	8.4	646	12 BI289102	BI289102 UI-R-DK0-
42	165.4	8.3	851	13 BU201022	BU201022 603952191
43	164.6	8.2	600	12 BI988528	BI988528 4012-24 M
44	158.2	7.9	569	29 CE228097	CE228097 tigr-gss-
45	158	7.9	405	14 CB811308	CB811308 AMGNNUC:S

ALIGNMENTS

RESULT 1	BC036103	3870 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	BC036103				
DEFINITION	Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (CDNA clone IMAGE:5274415), with apparent retained intron.				
ACCESSION	BC036103				
VERSION	BC036103.1				GI:23271310
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1. (bases 1 to 3870)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,				

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 12477932 2 (bases 1 to 3870) Strausberg,R. Direct Submission Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-re@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 48 Row: f Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein This clone has the following problem: retained intron.

FEATURES Location/Qualifiers 1..3870 /organism="Homo sapiens" /mol_type="mrna" /db_xref="taxon:9606" /clone="IMAGE:5274415" /tissue type="Brain, hippocampus" /clone_lib="NIH_MGC_95" /lab_host="DH10B" /note="Vector: pBluescript"

ORIGIN

Query Match 33.9%; Score 678.4; DB 11; Length 3870; Best Local Similarity 97.7%; Pred. No. 2.7e-142; Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0; QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60 DB 1143 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1202 QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 120 DB 1203 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 1262 QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180 DB 1263 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1322 QY 181 GCCCATCAGGCGGGTGGTAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240 DB 1323 GCCCATCAGGCGGGTGGTAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1382 QY 241 AAATTATCAGACATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAATTCAAGA 300 DB 1383 AAATTATCAGACATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAATTCAAGA 1442

QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 360 DB 1443 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 1502 QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420 DB 1503 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1562 QY 421 ACAAGGAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 480 DB 1563 ACAAGGAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 1622 QY 481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540 DB 1623 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 1682 QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 600 DB 1683 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 1742 QY 601 GAACAACCTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660 DB 1743 GAACAACCTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 1802 QY 661 TGGGTTCTTTTACAAGACCAGCCCTGACCTAGCTCCTGGACTGAC 704 DB 1803 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTAC 1846

RESULT 2

AY399453 5691 bp DNA linear GSS 12-DEC-2003 LOCUS Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY399453

VERSION AY399453.1 GI:39755442

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5691)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

Gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 5691)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source

1..5691 /organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>5691

/locus_tag="HCM0229"

gene

ORIGIN

Query Match 33.9%; Score 678.4; DB 29; Length 5691; Best Local Similarity 97.7%; Pred. No. 2.9e-142; Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db	794	GATTGGCAGAGAAAA	808		
RESULT 4					
AK044536					
LOCUS					
DEFINITION					
AK044536	3056 bp	mrna	linear	HTC 20-SEP-2003	
Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930019F21 product:dystrophin, muscular dystrophy, full insert sequence.					
AK044536					
AK044536.1	GI:26090404				
HTC; CAP trapper.					
Mus musculus (house mouse)					
SOURCE					
ORGANISM					
Mus musculus					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE					
1					
Carninci,P. and Hayashizaki,Y.					
High-efficiency full-length cDNA cloning					
Meth. Enzymol. 303, 19-44 (1999)					
99279253					
10349636					
REFERENCE					
2					
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.					
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
Genome Res. 10 (10), 1617-1630 (2000)					
20499374					
11042159					
REFERENCE					
3					
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.					
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer					
Genome Res. 10 (11), 1757-1771 (2000)					
20530913					
11076861					
REFERENCE					
4					
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.					
Functional annotation of a full-length mouse cDNA collection					
Nature 409, 685-690 (2001)					
5					
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.					
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs					
Nature 420, 563-573 (2002)					
6					
(bases 1 to 3056)					
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.					
Direct Submission					
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
A					

Query Match		24.6%;	Score 491.4;	DB 12;	Length 750;
Best Local Similarity		88.4%;	Pred. No. 3.4e-100;		
Matches	Conservative	0;	Mismatches	71;	Indels
					Gaps
					8;
					Gaps
					6;
QY	804	GATGTTGGAGFPACCTACTCATAGATTACTGTCAACAGATTCCCTCGACCTGGACCTGGAAAAAGTT	863		
DB	69	GCTGCTTGGGACAGAAACTCATAGATTACTGCAGAGTTCCTCTGGACCTGGAGAAGTT	128		
QY	864	TCTTGCCCTGG-CTTACAGAAAGCTGAAAC--AACTGCCAATGTCTCTACAGGATGCTACCCG	920		
DB	129	TCTTTCTCGATTACGGAAGCAGAAACAGAGCTGCGCAATGTCTCTACAGGACGCTTCCCG	188		
QY	921	-TAAGGAAAGGCTCCTAGAAAGACTCCAAGG--AGTAAAAAGAGCTGATGAAACAATGGCA	977		
DB	189	CTAAGGAGAAGCTCCTAGCAAGACTCCAGGGCAGTCAGAGAGCTGATGAAACCATGGCA	248		
QY	978	AGACCTCCAAGGTGAAA-TTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACA	1036		
DB	249	AGATCTCCAAGGAGAAAGTTGAAACTCACACAGATATCTATCACAATCTTGATGAAAAATG	308		
QY	1037	GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTT	1096		
DB	309	GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGGAAGCACCCCTGTTACAAAGACGTT	368		
QY	1097	TGGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCC	1156		
DB	369	TGGATAACATGAATTTCAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGGTCCC	428		
QY	1157	ATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTCTGG	1216		
DB	429	ATTTGGAAGCAAGTTCTGACCAGTGGAAAGCGTTTGATCTTTCTCTTCAGGAACCTCTTG	488		
QY	1217	TGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTC	1276		
DB	489	TTTGGCTACAGCTGAAAAGATGATGAACCTGAGCCCTCAGGCACCCATCGTGGTGATTTCC	548		
QY	1277	CAGCAGTTCAGAAAGCA- GAAAGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAA	1335		
DB	549	CAGCAGTTCAGAAAGCAGCAATGATATACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAA	608		
QY	1336	GAACTGTAACTATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTG	1395		
DB	609	GAACTGTAACTATGAGTACTCTTGAGACTGTGAGAATATTTCTGACAGAGCAGCCTTTG	668		
QY	1396	GAAGGACTAGAGAACTCTACCAGGAGCCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAG	1455		
DB	669	GAAGGACTAGAGAACTCTACCAGGAGCCCAGAGAACTGCCTCTCTGAAGAAAGAGCTCAG	728		
QY	1456	AATGTCACCTCGGTTCTTACGA	1476		
DB	729	AATGTCACCTCGGTTCTTACGA	749		

RESULT	7
AL121550	
LOCUS	AL121550
DEFINITION	DKFZp762L078_r1 762 (synonym: hmcl2) Homo sapiens cDNA clone DKFZp762L078_5', mRNA sequence.
	579 bp mRNA linear EST 04-SEP-2003

RESULT 8
CB228986
LOCUS
DEFINITION. CB228986 728 bp mrna linear EST 10-FEB-2003
AGENCOURT_11499247 NICHD_Rh_Ovl Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)

ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
1 (bases 1 to 728)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLC3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD Rh Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

FEATURES
source
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD Rh Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN
Query Match 23.6%; Score 471.6; DB 14; Length 728;
Best Local Similarity 95.3%; Pred. No. 9.9e-96;
Matches 486; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
195 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATGGAAACAGGAAATATCAGAAGA 254
1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATGGAAACAGGAAATATCAGAAGA 60
255 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCCTCAG 314
61 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCCTCAG 120
315 GGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCA 374
121 GGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCA 180
375 GAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAAGAAACAGGAAAAATGGA 434
181 GAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAAGAAACAGGAAAAATGGA 240
435 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 494
241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 300
495 GCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGT 554
301 GCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGT 360
555 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 614
361 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 420
615 ATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACA 674
421 ATTGGGAGATCGATGGGCAACATCTGCAGATGGACAGAACCGCTGGGTCTTTTACA 480
675 AGACCGCCTGACCTAGCTCCTGGACTGAC 704

Db 481 AGACATCCTTCTCAAATGGCAACGCTCTTAC 510
RESULT 9
CB850319
LOCUS CB850319 1047 bp mRNA linear EST 01-SEP-2003
DEFINITION MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION CB850319
VERSION CB850319.1 GI:34380806
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1047)
Yu, J., Farjo, R., MacNee, S.P., Baehr, W., Stambolian, D.E. and Swaroop, A.
TITLE Annotation and analysis of 10,000 expressed sequence tags from developing mouse eye and adult retina
JOURNAL Genome Biol. 4 (10), R65 (2003)
MEDLINE 22881944
PUBMED 14519200
COMMENT Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu.
Location/Qualifiers
1..1047
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="retina"
/clone_lib="MOUSE ADULT RETINA"
/note="Vector: pSPORT1"

FEATURES
source
1..1047
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="retina"
/clone_lib="MOUSE ADULT RETINA"
/note="Vector: pSPORT1"

ORIGIN
Query Match 23.0%; Score 459.4; DB 14; Length 1047;
Best Local Similarity 85.8%; Pred. No. 6.1e-93;
Matches 552; Conservative 0; Mismatches 87; Indels 4; Gaps 4;
33 GGACCGTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACAC 92
86 GGATAGTTACCAAACTGCTTTTANAGAAGTACTTTTCATGGCTTCTTCTGCCGAGGATAC 145
93 ATTGCAAGCACAAGGAGAGATTTTCTAATGATGTGGAAGTGGTGAAGACCAGTTTCATAC 152
146 ATTGGAGCACAAGGAGAGATTTTCAAAATGATGTTGAAGAAGTGAAGAAGACAGTTTCATGC 205
153 TCATGAGGGGTACATGATGGATTTCAGACGCCCATCAGGGCGGGTTGGTAATATTCTACA 212
206 TCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTGGTAATGTTCTACA 265
213 ATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAATGAAGAACTGAAGTACA 272
266 GTTAGGAAGTCAACTAGTTGGAAAAGGGAAAAATTATCAGAAATGAAGAAGCTGAAGTGCA 325
273 AGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAG-CTAGCATGGAAA 331
326 AGAACAAATGAATCTCTAAATTCAGATGGGAATGTCCTCAGGGTANCCCTAGCATGGAAA 385
332 AACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAACTGAAAAGAGTTGA 391
386 AACAAAGCAATTCACAAAGTTCTTAATGGATCTCCAGAAATCAGAAATTAAGAAGACTAG 445
392 ATGACTGGCTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGAGCCCTCTTGGAC 451
446 ATGACTGGTTAACAAAAACTGAAGAGAGAACTAANAATAATGGAGGAAGAGCCCTTTGGAC 505
452 CTGATCTTGAAGACCTAAACGCCCAAGTACAAACATAAAGTGC-TTCAAGAAGATCTA 510

Db 506 CTGATCTTGAAGATCTAAATGCCAAGTACAACAACATAAGGTGCTTTCAAGAAGATCTA 565

QY 511 GAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCT 570

Db 566 GAACAGGAGCAGGTTCAGGGTCAACTCGCTCACTCACATGGNAGTAGTGGGTGATGAATCC 625

QY 571 AGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGGAGATCGATG 629

Db 626 AGCGGGGATCATGCAACAGCTGCTTTTGGGAANAACAACCTTAAGGNACTGGGAGATCGATG 685

QY 630 GGCAACATCTG-TAGATGGACAGAACGCGCTGGGTTCTTTT 671

Db 686 GGCAATATCTGCAAAATGGACTGAAACCCGTTGGATTGTTTT 728

RESULT 10

AY399454

LOCUS

AY399454 5676 bp DNA linear GSS 12-DEC-2003

DEFINITION

Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION

AY399454

VERSION

AY399454.1 GI:39755443

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1 (bases 1 to 5676)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 5676)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering them based on alignment.

Location/Qualifiers

1. .5676

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..>5676

/locus_tag="HCM0229"

ORIGIN

Query Match 22.3%; Score 446; DB 29; Length 5676;

Best Local Similarity 84.0%; Pred. No. 9.4e-90;

Matches 458; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 160 GGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGGTAATATCTACAATTGGGA 219

Db 529 GGGTACATGATGGATNNNNNAGCCCATCAGGGCCGGTGGTAATANNNNNNANTTGGGA 588

QY 220 AGTAAGCTGATTGGAACAGGAAAATTATCAGACATGAAGAACTGAAGTACAAGAGCAG 279

Db 589 AGTNNNTGATTGGAACAGGAAAATTATCAGACATGAAGAACTGAAGTACAAGAGCAG 648

QY 280 ATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGCTAGCATGGAACAAAGC 339

Db 649 ATGAATCTNNNNATTCAGATGGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 708

QY 340 AATTACATAGATTTTAAATGGATCTCCAGAACTCAGAAACTGAAGAGTTGAATGACTGG 399

Db 709 NNTTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGG 768

QY 400 CTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTT 459

Db 769 CTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTT 828

QY 460 GAAGACCTAAAAACGCCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAA 519

Db 829 GAAGACCTAAAAACGCCCAAGTACAACAACATAAGNNNNNNNCAAGAAGATTTAGAACAAAGAA 888

QY 520 CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 579

Db 889 CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 948

QY 580 CACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATC 639

Db 949 CACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATC 1008

QY 640 TGTAGATGGACAGAACGCGCTGGTCTTTTACAAGACCAGCCTGACCTAGCTCCTCGGA 699

Db 1009 TGTAGTGGACAGAACGCCGCTGGTCTTTTACAAGACATCCTTCTCAATGCGCAACGT 1068

QY 700 CTGAC 704

Db 1069 CTTAC 1073

RESULT 11

BM488464

LOCUS

DEFINITION

BM488464 663 bp mRNA linear EST 07-FEB-2002

pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA clone pgm2n.pk007.117 5' similar to sp|P11533|DMD CHICK Dystrophin p1r|S02041 dystrophin, muscle - chicken emb|CAA31746.1| (X13369) dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.

BM488464

ACCESSION

BM488464.1 GI:18609395

VERSION

EST.

KEYWORDS

Gallus gallus (chicken)

SOURCE

Gallus gallus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 663)

Cogburn,L.A. and Monsonogo-Ornan,E.

ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome Project

Unpublished (2002)

Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

Location/Qualifiers

1. .663

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler and Ottawa Res. Centre Strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk007.117"

/sex="Male and Female"

/tissue type="Breast muscle, leg muscle and epiphyseal growth plate"

/dev stage="Breast, leg:Embryo(d19);post-hatch(1d,1,3,5,7,9,11 weeks);growth plate(1d,7d,14d post-hatch)"

/lab_host="E. coli EMDH108"

/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"

/note="Vector: pCMVSPORT6; Library made from equivalent source

FEATURES

source

pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

ORIGIN

Query Match	21.2%;	Score 424.2;	DB 12;	Length 663;
Best Local Similarity	78.1%;	Pred. No. 5e-85;		
Matches 518;	Conservative 0;	Mismatches 144;	Indels 1;	Gaps 1;
Qy	960	GCTGATGAAACAATGGCAAGACCTCCAAAGGTGAATGAAGCTCACACAGATGTTTATCA	1019	
Db	1	GCTCATGAAGCAGTGGCAGGATCTACAGGCAGAAATTGATGCACATACTGACATCT-TNN	59	
Qy	1020	CAACCTGGATGAAACACGCCAAATAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGT	1079	
Db	60	CAACCTGGATGAAACACGCCAGAAATCCTGAGATCCCTGGAAAGTTCAGAGGATGCTGT	119	
Qy	1080	CCTGTTACAAAGACGTTTGGATAACATGAATCAAGTGGAGTGAAGTTCGGAAAAAGTC	1139	
Db	120	CCTGTTGCAGAGACGCTCGGATAACATGAATCAAGTGGAGTGAAGTTCAGAGGATG	179	
Qy	1140	TCTCAACATTAGTCCCAATTTGGAAGCCAGTTCGACCATGGAAGCGTCTGCACCTTTC	1199	
Db	180	TCTAAACATTAGATCTCAATTTGGAAGCCAGCACAGACCATGGAAGCGTTTACATCTCTC	239	
Qy	1200	TCTGCAGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC	1259	
Db	240	TCTTCAGGAACCTTTTGGCATGGCTGCAATTTGAAGGAGGATGAATTAAGCAGCAAGCACC	299	
Qy	1260	TATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAG	1319	
Db	300	CATTGGTGAGATATTCCCACTGTGCAGAAAGCAGAAATGATGTTTCATAGGACTTTTCAAGAG	359	
Qy	1320	GGAAATTGAAACTAAAGAACCTGTAAATCATGACTACTCTTGAGACTGTACGAATATTTCT	1379	
Db	360	GGAGCTGAAACAAAGAACCTGTTATCATGATGCACTTGAGACTGTGCGACTCTTCTCT	419	
Qy	1380	GACAGAGCAGCCTTTTGGAGGACTAGAGAAACTCTACCAGGAGCCCAAGAGAGCTGCCTCC	1439	
Db	420	GGCAGATCAACCCAGTAGAGGACTGGAAAGGTCTATCCAGAACCAAGACAGACCTATCACC	479	
Qy	1440	TGAGAGAGAGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC	1499	
Db	480	TGAGAGAGGGCCCGAGATGTCACTAAAGTTCTCGAAGGCAAGCAGATGATGTCAGNAC	539	
Qy	1500	TGAGTGGGAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTT	1559	
Db	540	TGAGTGGGATAAGCTAAATCTACGTTCTGCTGATGGCAAAAGAGATAGATGATGCTCT	599	
Qy	1560	TGAAAGACTCCAGGAACITCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGC	1619	
Db	600	TGAAAGACTGCAGGGTCTTCAGGAGGCAATGGATGAACCTNNNCCTGAAACTGCGCCAGGC	659	
Qy	1620	TGA 1622		
Db	660	TGA 662		

RESULT 12
CB177816/c
LOCUS CB177816 595 bp mRNA linear EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3' similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
CB177816
ACCESSION CB177816.1 GI:28186206
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

TITLE
JOURNAL
COMMENT

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,B., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.

FEATURES
source

1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match	20.4%;	Score 409;	DB 14;	Length 595;
Best Local Similarity	100.0%;	Pred. No. 1.3e-81;		
Matches 409;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA	60	
Db	409	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA	350	
QY	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT	120	
Db	349	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT	290	
QY	121	GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	180	
Db	289	GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	230	
QY	181	GCCCATCAGGCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA	240	
Db	229	GCCCATCAGGCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA	170	
QY	241	AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGACATGATCTCCTAAATCAAGA	300	
Db	169	AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGACATGATCTCCTAAATCAAGA	110	
QY	301	TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG	360	
Db	109	TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG	50	
QY	361	GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAA 409		
Db	49	GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAA 1		

Db 402 TGATGAGTTGAGCGGCGAGGCACCTATTGGTGGCGATTTTCCAGCAGTTTCAGAAAGCAGAA 343
QY 1296 CGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTAAATCATGAGTAC 1355
Db 342 TGATGTACACAGGCGCTTCAAGAGGGGAATTGANAACCTAAAGAACCTGTAAATCATGAGTAC 283
QY 1356 TCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTA 1415
Db 282 TCTGGAGACTGTGAGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTA 223
QY 1416 CCAGGAGCCCGACAGAGCTGCCTCCTGAGAGAGAGCCCGAGAATGTCACTCGGCTTCTACG 1475
Db 222 CCAGGAGCCCGACAGAGCTGCCTCCTGGAAGAAAGAGCTCAGAATGTCACTCGGCTCCTACG 163
QY 1476 AAAGCAGGCTGAGAGGTTCAATACTAGTGGGAAAATTGAACCTGCACCTCGGCTGACTG 1535
Db 162 AAAGCAGGCTGAGAGGTTCAACACTGAATGGGACAAATTGAACCTGCACCTCAGCTGATTG 103
QY 1536 GCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGA 1595
Db 102 GCAGCGAAAATAGATGAGGCTCTTGAAGACTGCAGGAACCTTCAAGAGCTGCGGATGA 43
QY 1596 GCTGGACCTCAAGCTGGCC 1615
Db 42 ACAGCACCTCATGTTGGGC 23

RESULT 15
AY408546
LOCUS
DEFINITION Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY408546 genomic survey sequence.
ACCESSION AY408546
VERSION AY408546.1 GI:39764517
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 2874)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2874)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2874
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>2874
/gene="DRP2"
/locus_tag="HCM3260"

ORIGIN

Query Match 15.5%; Score 310.8; DB 29; Length 2874;
Best Local Similarity 59.5%; Pred.No. 2.8e-59;
Matches 525; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

QY 1104 CATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGA 1163
Db 234 CATGAATCTGTGTGGAATGAATAAAAAAGAAGTCTCACAACCTCGCGCTCGCCTAGA 293
QY 1164 AGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAATTTCTGGTGTGGCT 1223
Db 294 GGCCTTCTCAGACCAACAGTGGAAAGCTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCT 353
QY 1224 ACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGT 1283
Db 354 CAGCCAAAAGGATGAGAGTTGTTCAGCTCAGCTGCCCTACAGGGGATGTGGCCCTGGT 413
QY 1284 TCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGT 1343
Db 414 GAAACAGGAGAAAGGACACATCGGCCCTTATGGAAGAAAGTCAAGTCTCGGGGCCCTTA 473
QY 1344 AATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGACT 1403
Db 474 CATCTATTCTGTGTGAGTCACTCAGGCTCAGCTCAGGCTTCTGTCCAGGACCCATTGAGGAGTT 533
QY 1404 AGAGAAACTCTACAGGAGCCCGACAGAGAGCTGCCCTCCTGAGGAGAGAGCCCGAGAATGTAC 1463
Db 534 AGAGGAGCCTCATTTCTGAGAGCAAGATACCTCCCCGAAACAGCGGATCCAGAAATCTCAG 593
QY 1464 TCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCA 1523
Db 594 CCGCTTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACGTGTGGGAGAAATTGACAGCCCG 653
QY 1524 CTCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTTCAGGAACTTCAAGA 1583
Db 654 CTGTGTGGACCAAGCACCGTCACATTTGAGCGGACTCTGGAGCAGCTCTTGGAGATTTCAAG 713
QY 1584 GGCACCGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCA 1643
Db 714 GGCAATGGAGGAACCTAAGCACTACTCTGAGCAAGCTGAGGGAGTCCGAGCCACTTGGGA 773
QY 1644 GCCGTGGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAAAGTCAAGGCACT 1703
Db 774 GCCCATTTGGGATCTCTTCAATTTGATTCACTCCAGAGCACATCCAGGCTATTAAAGCTGTT 833
QY 1704 TCGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACCGTCAATGACCTTGTCTGCCA 1763
Db 834 CAAAGAAGAAATTTCTCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCAACA 893
QY 1764 GCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAA 1823
Db 894 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAAATTCAGGCGCTTGAACAGATCAA 953
QY 1824 CACCAGATGGAAGCTTCTGCAAGTGGCCCTGAGGACCGAGTCAGGAGCTGCATGAAGC 1883
Db 954 CGTCGATGGAAACAACACTACAGGCGTCAGTTAGTGAGAGGCTTAAGCAGCTCCAGGATGC 1013
QY 1884 CCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTTCTGTCCAGGTTCCCTG 1943
Db 1014 CCACGGGACTTTGGGCTGGGTCAACAGCACTTTCTCTCTCTCTGTCCAGGTTCCCTG 1073
QY 1944 GGAGAGAGCCATCTCGCAAAACAAGTGGCCCTACTATATCAA 1985
Db 1074 GGAAGAGCAATTTTCAACCAATAAAGTTCCCTACTACATCAA 1115

Search completed: April 5, 2004, 08:15:04
Job time : 3501.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 12:04:07 ; Search time 6259.59 Seconds
(without alignments)
16625.161 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 acatgcattcaacatgcga.....tggatctcttgagcctgtca 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1839.4	76.6	5462	6	AX538621	AX538621 Sequence
2	1725	71.8	5339	6	AX538620	AX538620 Sequence
3	1419	59.1	5417	6	AX538619	AX538619 Sequence
4	1335.6	55.6	5952	6	AR304538	AR304538 Sequence
5	1335.6	55.6	5952	6	AX114289	AX114289 Sequence
6	1335.6	55.6	8689	6	AX538622	AX538622 Sequence
7	1335.6	55.6	11443	6	AX538624	AX538624 Sequence
8	1335.6	55.6	12057	6	AX538627	AX538627 Sequence
9	1335.6	55.6	12446	9	HSDMDR	X14298 Human mRNA
10	1335.6	55.6	13957	6	AX409637	AX409637 Sequence
11	1335.6	55.6	13957	6	AX538581	AX538581 Sequence
12	1335.6	55.6	13957	9	HUMDYS	M18533 Homo sapien
13	1335.6	55.6	13977	6	AR220819	AR220819 Sequence
14	1179.4	49.1	13887	4	AF070485	AF070485 Canis fam
15	1143.6	47.6	13815	6	AX306153	AX306153 Sequence
16	1143.6	47.6	13815	6	AX538582	AX538582 Sequence
17	1143.6	47.6	13815	10	MUSDYSA	M68859 Mouse dystro
18	1143.6	47.6	19307	6	AR093392	AR093392 Sequence
19	1143.6	47.6	19307	6	AR142592	AR142592 Sequence
20	1093.8	45.6	4402	6	E30220	E30220 Shortened d
21	1021.2	42.5	4402	6	E30219	E30219 Shortened d
22	1018.8	42.4	4075	6	E30221	E30221 Shortened d
23	879.6	36.6	13575	5	GGDYS	X13369 Chicken mRN
24	814.6	33.9	1481	9	BC009242	BC009242 Homo sapi
25	709.2	29.5	3747	6	E30218	E30218 Shortened d
26	644.4	26.8	3275	10	MUSDYS	M18025 Mouse dystro
27	530.2	22.1	3521	5	AF339031	AF339031 Danio rer
28	498.6	20.8	10705	10	RNAJ2967	AJ002967 Rattus no
29	495.6	20.6	11096	6	AX538584	AX538584 Sequence
30	495.6	20.6	11096	10	MMY12229	Y12229 M.musculus
31	492.2	20.5	2654	5	FSCDYSTRO	M37645 Torpedo cal
32	487	20.3	6045	6	A63605	A63605 Sequence 7
33	487	20.3	6045	6	AR281528	AR281528 Sequence
34	487	20.3	6059	6	AX107972	AX107972 Sequence
35	487	20.3	10302	6	AX538583	AX538583 Sequence
36	487	20.3	10302	9	HSMUPS	X69086 H.sapiens m
37	487	20.3	10320	6	A63607	A63607 Sequence 9
38	487	20.3	10320	6	AR281529	AR281529 Sequence
39	476.4	19.8	630	9	HSDMDF1	X06179 Human fetal
40	459	19.1	3161	10	MMGUTRPH	X83506 M.musculus
41	411.8	17.2	3499	9	HSU43519	U43519 Human dystro
42	411.8	17.2	5106	6	AX552248	AX552248 Sequence
43	403.8	16.8	3073	10	AF195788	AF195788 Rattus no
44	403.8	16.8	3140	10	AF195787	AF195787 Rattus no
45	395.8	16.5	4416	10	BC060605	BC060605 Mus muscu

ALIGNMENTS

RESULT 1					
AX538621					
LOCUS	AX538621	5462 bp	DNA	linear	PAT 23-NOV-2002
DEFINITION	Sequence 41 from Patent WO229056.				
ACCESSION	AX538621				
VERSION	AX538621.1	GI:25271168			
KEYWORDS					
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
	artificial sequences.				
REFERENCE	1				
AUTHORS	Chamberlain,J.S. and Harper,S.Q.				
TITLE	Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL	Patent: WO 0229056-A 41 11-APR-2002;				
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

FEATURES
Source Location/Qualifiers
1..5462
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN

Query Match 76.6%; Score 1839.4; DB 6; Length 5462;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 76; Indels 210; Gaps 2;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
DB 799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 858
QY 61 TGTGTATPACCACCTATCCAGATPAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
DB 859 TGTGTATPACCACCTATCCAGATPAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 918
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTGCCAAGGCC 180
DB 919 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTGCCAAGGCC 978
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 240
DB 979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 1038
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
DB 1039 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1098
QY 301 GAGCTATGCTTACACACAGGCTGCTTATGTCAACCACCTGTGACCCCTACACGAGCCCAT 360
DB 1099 GAGCTATGCTTACACACAGGCTGCTTATGTCAACCACCTGTGACCCCTACACGAGCCCAT 1158
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATGTATGGAGAG 420
DB 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATGTATGGAGAG 1218
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
DB 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 481 TGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 540
DB 1279 TGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338
QY 541 CCAGTTTCACTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 600
DB 1339 CCAGTTTCACTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 601 TAATATTCTAACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTATCAGAAATGAAGA 660
DB 1399 TAATATTCTAACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTATCAGAAATGAAGA 1458
QY 661 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
DB 1459 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
DB 1519 TAGCATGGAAAAACAAAGCAATTTACATGCTCTCTGGACTG----- 1558
QY 781 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAGGAAAAATGGAGGAAGA 840
DB 1559 -----AC 1560
QY 841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAAACATAAGGTGCTTCA 900
DB 1561 CACTATTGGAGCCTCTCCTACTCAGACTGTACTCTGTGTGACACAACCTGTGGTTACTAA 1620
QY 901 AGAAGATCTAGAACAAGAAACAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGGTAGT 960

DB 1621 GGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGATGTTGGAG----- 1669
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
DB 1670 ----- 1669
QY 1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAC 1080
DB 1670 ----- 1669
QY 1081 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA 1140
DB 1670 -CATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA 1728
QY 1141 AGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCGTTAAGGAAAGGCTCCTAGAAGA 1200
DB 1729 AGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCGTTAAGGAAAGGCTCCTAGAAGA 1788
QY 1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGC 1260
DB 1789 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGC 1848
QY 1261 TCACACAGATGTTTATCACAACTCTGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGA 1320
DB 1849 TCACACAGATGTTTATCACAACTCTGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGA 1908
QY 1321 AGTTCCGATGATGCAGTCTCTCAACATTAGTCCCATTTTGGATAACATGAACCTCAAGTGGAG 1380
DB 1909 AGTTCCGATGATGCAGTCTCTCAACATTAGTCCCATTTTGGATAACATGAACCTCAAGTGGAG 1968
QY 1381 TGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 1440
DB 1969 TGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 2028
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGCTACAGCTGAAAGATGATGA 1500
DB 2029 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGCTACAGCTGAAAGATGATGA 2088
QY 1501 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTTCCAGAGCAGAACCGATGT 1560
DB 2089 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTTCCAGAGCAGAACCGATGT 2148
QY 1561 ACATAGGGCCTTCAAGAGGGGAATTTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
DB 2149 ACATAGGGCCTTCAAGAGGGGAATTTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2208
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 1680
DB 2209 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 2268
QY 1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
DB 2269 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 2328
QY 1741 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAAACCTGCACTCCGCTGACTGGCAGAG 1800
DB 2329 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAAACCTGCACTCCGCTGACTGGCAGAG 2388
QY 1801 AAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGCCCAAGGATGAGTGGGA 1860
DB 2389 AAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGCCCAAGGATGAGTGGGA 2448
QY 1861 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGACGCCCTGGCGGATCTCCT 1920
DB 2449 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGACGCCCTGGCGGATCTCCT 2508
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGGCC 1980
DB 2509 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGGCC 2568
QY 1981 TCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 2040

Db 2569 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2628

QY 2041 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAAACACCCAGATGGAAGCTTCT 2100

Db 2629 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAAACACCCAGATGGAAGCTTCT 2688

QY 2101 GCAGGTGGCCGTCGAGGACCCGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTGGTCC 2160

Db 2689 GCAGGTGGCCGTCGAGGACCCGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTGGTCC 2748

QY 2161 AGCATCTCAGCACTTCTTTCCAGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCC 2220

Db 2749 AGCATCTCAGCACTTCTTTCCAGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCC 2808

QY 2221 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGCTGGGACCATCCCAA 2280

Db 2809 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGCTGGGACCATCCCAA 2868

QY 2281 AATGACAGAGCTCTACCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340

Db 2869 AATGACAGAGCTCTACCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2928

QY 2341 GACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400

Db 2929 GACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2988

QY 2401 A 2401

Db 2989 A 2989

RESULT 2

AX538620 5339 bp DNA linear PAT 23-NOV-2002

LOCUS

DEFINITION Sequence 40 from Patent WO0229056.

ACCESSION AX538620

VERSION AX538620.1 GI:25271166

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS Chamberlain, J.S. and Harper, S.Q.

TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229056-A 40 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

source

1..5339

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Synthetic"

ORIGIN

Query Match 71.8%; Score 1725; DB 6; Length 5339;

Best Local Similarity 86.1%; Pred. No. 0;

Matches 2068; Conservative 0; Mismatches 0; Indels 333; Gaps 1;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCCTGAAGA 60

Db 799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCCTGAAGA 858

QY 61 TGTGTATACCACCTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120

Db 859 TGTGTATACCACCTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 918

QY 121 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180

Db 919 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 978

QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATATGCACTATTCTCAACA 240

Db 979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATATGCACTATTCTCAACA 1038

QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 300

Db 1039 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 1098

QY 301 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCTCTACACGGAGCCCAT 360

Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCTCTACACGGAGCCCAT 1158

QY 361 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420

Db 1159 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 480

Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATATCGTGGCTTCTTTC 1278

QY 481 TGTGTAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 540

Db 1279 TGTGTAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 1338

QY 541 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 600

Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1398

QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 660

Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1458

QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 720

Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518

QY 721 TAGCATGGAAAAACAACCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780

Db 1519 TAGCATGGAAAAACAACCAATTTTACA----- 1545

QY 781 GAAAGAGTTGAATGACTCGCTAACAAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA 840

Db 1546 ----- 1545

QY 841 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 900

Db 1546 ----- 1545

QY 901 AGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 960

Db 1546 ----- 1545

QY 961 TGATGAATCTAGTGGAGATCAGCCTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020

Db 1546 ----- 1545

QY 1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC 1080

Db 1546 ----- 1545

QY 1081 TCATAGATTACTGCAACAGTTCCCTCGACCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1140

Db 1546 TCATAGATTACTGCAACAGTTCCCTCGACCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1605

QY 1141 AGCTGAAACAACTGCCAATGTCTCTACAGGATGTCTACCCGTAAAGGAAAGGCTCCTAGAAGA 1200

Db 1606 AGCTGAAACAACTGCCAATGTCTCTACAGGATGTCTACCCGTAAAGGAAAGGCTCCTAGAAGA 1665

QY 1201 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1260

Db 1666 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1725

QY 1261 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320

Db 1726 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1785

QY 1321 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACATTCAAAGTGGAG 1380
Db |||||
QY 1786 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACATTCAAAGTGGAG 1845
Db |||||
QY 1381 TGAAC TTCGAAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCAGTG 1440
Db |||||
QY 1846 TGAAC TTCGAAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCAGTG 1905
Db |||||
QY 1441 GAAGCGTCTGCACCTTTCTCTGCGAGGAACCTTCTGGTGTGGGTACAGCTGAAGATGATGA 1500
Db |||||
QY 1906 GAAGCGTCTGCACCTTTCTCTGCGAGGAACCTTCTGGTGTGGGTACAGCTGAAGATGATGA 1965
Db |||||
QY 1501 ATTAAGCCGCGAGGCACCTATTGGAGCGCACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 1560
Db |||||
QY 1966 ATTAAGCCGCGAGGCACCTATTGGAGCGCACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 2025
Db |||||
QY 1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGA 1620
Db |||||
QY 2026 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGA 2085
Db |||||
QY 1621 GACTGTACGAATATTCTGACAGAGCGACCTTTTGAAGGACTAGAGAACTCTACCAAGGA 1680
Db |||||
QY 2086 GACTGTACGAATATTCTGACAGAGCGACCTTTTGAAGGACTAGAGAACTCTACCAAGGA 2145
Db |||||
QY 1681 GCCCAGAGAGCTGCCTCCTGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db |||||
QY 2146 GCCCAGAGAGCTGCCTCCTGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2205
Db |||||
QY 1741 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
Db |||||
QY 2206 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2265
Db |||||
QY 1801 AAAATAGATGACACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1860
Db |||||
QY 2266 AAAATAGATGACACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 2325
Db |||||
QY 1861 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1920
Db |||||
QY 2326 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 2385
Db |||||
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTTGCGCC 1980
Db |||||
QY 2386 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTTGCGCC 2445
Db |||||
QY 1981 TCTGAAAGAGACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 2040
Db |||||
QY 2446 TCTGAAAGAGACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 2505
Db |||||
QY 2041 TCAGCTCTCACCGTATAACCTCAGCCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT 2100
Db |||||
QY 2506 TCAGCTCTCACCGTATAACCTCAGCCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT 2565
Db |||||
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCA CAGGACCTTTGGTCC 2160
Db |||||
QY 2566 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCA CAGGACCTTTGGTCC 2625
Db |||||
QY 2161 AGCATCTCAGCCTCTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
Db |||||
QY 2626 AGCATCTCAGCCTCTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2685
Db |||||
QY 2221 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACACTTGCTGGGACCATCCCAA 2280
Db |||||
QY 2686 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACACTTGCTGGGACCATCCCAA 2745
Db |||||
QY 2281 AATGACAGAGCTCTACAGAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAG 2340
Db |||||
QY 2746 AATGACAGAGCTCTACAGAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAG 2805
Db |||||
QY 2341 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db |||||
QY 2806 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGCTTGGATCTCTTGAGCCTGTC 2865
Db |||||
QY 2401 A 2401

Db 2866 A 2866
RESULT 3
AX538619
LOCUS AX538619
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 39 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1. 5417
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"
ORIGIN
Query Match 59.1%; Score 1419; DB 6; Length 5417;
Best Local Similarity 78.1%; Pred.No. 0;
Matches 1876; Conservative 0; Mismatches 270; Indels 255; Gaps 4;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db |||||
QY 799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 858
Db |||||
QY 61 TGTTGATACCACTATCCAGATAAGAGTCCATCTTAATGTATCATCATCACTCTTCCA 120
Db |||||
QY 859 TGTTGATACCACTATCCAGATAAGAGTCCATCTTAATGTATCATCATCACTCTTCCA 918
Db |||||
QY 121 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC 180
Db |||||
QY 919 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC 978
Db |||||
QY 181 ACCTAAAGTGACTAAAGAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 240
Db |||||
QY 979 ACCTAAAGTGACTAAAGAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 1038
Db |||||
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 300
Db |||||
QY 1039 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 1098
Db |||||
QY 301 GAGTATGCCCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 360
Db |||||
QY 1099 GAGTATGCCCTACACACAGGCTGCTTATGTACCACTCTGTACCCCTACACGGAGCCCAT 1158
Db |||||
QY 361 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCAATTGATGGAGAG 420
Db |||||
QY 1159 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCAATTGATGGAGAG 1218
Db |||||
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
Db |||||
QY 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
Db |||||
QY 481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db |||||
QY 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
Db |||||
QY 541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 600
Db |||||
QY 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1398
Db |||||
QY 601 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db |||||

Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATATCAGAAGATGAAGA 1458
Qy 661 AACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 1459 AACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
Qy 721 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 1519 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
Qy 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 840
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1638
Qy 841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 900
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 1698
Qy 901 AGAAGATCTAGAACAGAAACAAAGTCAGGCTCAATTTCTCACTCACATGGTGGTAGT 960
Db 1699 AGAAGATCTAGAACAGAAACAAAGTCAGGCTCAATTTCTCACTCACATGGTGGTAGT 1758
Qy 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
Qy 1021 AGATCGATGGCAACACATCTGTAGATGGACAGAAAGCCGCTGGGTCTTTTACAAGACAC 1080
Db 1819 AGATCGATGGCAACACATCTGTAGATGGACAGAAAGCCGCTGGGTCTTTTACAAGACAT 1878
Qy 1081 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAGTTTCTTGCTGCTTACAGA 1140
Db 1879 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1938
Qy 1141 AGCTGAAACAACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAGA 1200
Db 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTA-----AAGA 1983
Qy 1201 CTCCAAGGGAGTAAAGAGCTGATGAACAACATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
Db 1984 TCAAAATGAAATGTTATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAA 2043
Qy 1261 TCACACAGATGTTTATCAAACTGGATGAAACAGCCAAATAATCCTGAGATCCCTGGA 1320
Db 2044 GAAAAGCAATCCATGGGCAAACTGTATTCACTCAACAAGATTTCTTTCAACACTGAA 2103
Qy 1321 AGGTTCCGATGATGCAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 1440
Db 2104 GAATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGA 2163
Qy 1381 TGAACCTTCGGAAGAAAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 1440
Db 2164 TAATTTAGTCCAAAAGTGTGAAAAGAGTACAGCACA----- 2199
Qy 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGCTGGCTACAGCTGAAAGATGATGA 1500
Db 2200 ----- 2199
Qy 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1560
Db 2200 -----GATTTACAGGCTGTCAACCACCACTCAGCCAT 2231
Qy 1561 ACATAGGCGCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTTTGA 1620
Db 2232 CACTAAACACAGACAACTGTAATGGAACAGTAACACTACGGTGACCAAGGGAACA----- 2286
Qy 1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAGGA 1680
Db 2287 ----- 2286
Qy 1681 GCCCAGAGAGCTGCCTCTGTAGGAGAGAGGCCAGAGATGTCACTCGGCTTCTACGAAAGCA 1740
Db 2287 -----GATCCTGGTAAAGCA 2301

Qy 1741 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1800
Db 2302 TGCTCAAGAGG-----AACTTCCACCACCACCTCCCAAAAAGAGAG 2343
Qy 1801 AAAATATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGATGAGCTGGA 1860
Db 2344 GCAGATTACTGTGGATCTTGAAGACTCCAGGAACTTCAAGAGGCCACGATGAGCTGGA 2403
Qy 1861 CCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCCTGGGCGATCTCCT 1920
Db 2404 CCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCCTGGGCGATCTCCT 2463
Qy 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC 1980
Db 2464 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC 2523
Qy 1981 TCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 2040
Db 2524 TCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 2583
Qy 2041 TCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
Db 2584 TCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2643
Qy 2101 GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCC 2160
Db 2644 GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCC 2703
Qy 2161 AGCATCTCAGCACCTTTCTTCCACGTCGTCTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2220
Db 2704 AGCATCTCAGCACCTTTCTTCCACGTCGTCTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2763
Qy 2221 AAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAATTGCTGGGAGCCATCCCAA 2280
Db 2764 AAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAATTGCTGGGAGCCATCCCAA 2823
Qy 2281 AATGACAGAGCTCTACAGTCTTTAGTGCCTGAAATAATGTCAAGATTCAGCTTATAG 2340
Db 2824 AATGACAGAGCTCTACAGTCTTTAGTGCCTGAAATAATGTCAAGATTCAGCTTATAG 2883
Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db 2884 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2943
Qy 2401 A 2401
Db 2944 A 2944

RESULT 4
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. 5952
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 55.6%; Score 1335.6; DB 6; Length 5952;
Best Local Similarity 98.3%; Pred. No. 0;

3142 GATGTTTATCAACACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCC 3201
1328 GATGATGCACTCCTGTTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 1387
3202 GATGATGCACTCCTGTTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 3261
1388 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCCAGTGGAGCGGT 1447
3262 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCCAGTGGAGCGGT 3321
1448 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
3322 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 3381
1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAACGACGAACGATGTACATAGG 1567
3382 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAACGACGAACGATGTACATAGG 3441
1568 GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTA 1627
3442 GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTA 3501
1628 CGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCCAGGAGCCCGAGA 1687
3502 CGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCCAGGAGCCCGAGA 3561
1688 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
3562 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 3621
1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 1807
3622 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 3681
1808 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
3682 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 3741
1868 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 1927
3742 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 3801
1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCCTCTGAAA 1987
3802 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCCTCTGAAA 3861
1988 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 2047
3862 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 3921
2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTTCGAGGTG 2107
3922 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTTCGAGGTG 3981
2108 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCT 2167
3982 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCT 4041
2168 CAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCCATCTCGCCCAACAAA 2227
4042 CAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCCATCTCGCCCAACAAA 4101
2228 GTGCCCTACTATATCAACACGAGACTCAAACTTGTGGGACCATCCCAAAATGACA 2287
4102 GTGCCCTACTATATCAACACGAGACTCAAACTTGTGGGACCATCCCAAAATGACA 4161
2288 GAGCTCTACCACTTTTAGCTGACCTGAATAATGTGAGTCTCAGCTTATAGGACTGCC 2347
4162 GAGCTCTACCACTTTTAGCTGACCTGAATAATGTGAGTCTCAGCTTATAGGACTGCC 4221
2348 ATGAAACTCCGAGACTGCAGAGGCGCCTTTGCTTGGATCTCTGAGCCTGTCA 2401
4222 ATGAAACTCCGAGACTGCAGAGGCGCCTTTGCTTGGATCTCTGAGCCTGTCA 4275

RESULT 6
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..8689
source /organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 55.6%; Score 1335.6; DB 6; Length 8689;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1028 TGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
Db 2948 TGGAGAAGCATTCATAAAAGGTGAGTGAGCGGAGAGGCTGCTTTTGGAAAGAACTCATAGA 3007
QY 1088 TTACTGCAACAGTTCCTTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 1147
Db 3008 TTACTGCAACAGTTCCTTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 3067
QY 1148 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 1207
Db 3068 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 3127
QY 1208 GGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 1267
Db 3128 GGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 3187
QY 1268 GATGTTTATCAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
Db 3188 GATGTTTATCAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 3247
QY 1328 GATGATGCAGTCTCTTACAAAGACGTTTGGATAAATCAATGAACCTTCAAGTGGAGTGAACCTT 1387
Db 3248 GATGATGCAGTCTCTTACAAAGACGTTTGGATAAATCAATGAACCTTCAAGTGGAGTGAACCTT 3307
QY 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCCAGTGGAGCGGT 1447
Db 3308 CGGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCCAGTGGAGCGGT 3367
QY 1448 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
Db 3368 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 3427
QY 1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG 1567
Db 3428 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG 3487
QY 1568 GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1627
Db 3488 GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 3547
QY 1628 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTTACAGGAGCCCGAGA 1687
Db 3548 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTTACAGGAGCCCGAGA 3607

QY	1688	GAGCTGCCTCCTGAGGAGAGAGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG	1747
DB	3608	GAGCTGCCTCCTGAGGAGAGAGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG	3667
QY	1748	GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATA	1807
DB	3668	GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATA	3727
QY	1808	GATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	1867
DB	3728	GATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	3787
QY	1868	CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC	1927
DB	3788	CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC	3847
QY	1928	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCCACTTCGAGGAGAAATTGCGCCTCTGAAA	1987
DB	3848	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCCACTTCGAGGAGAAATTGCGCCTCTGAAA	3907
QY	1988	GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC	2047
DB	3908	GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC	3967
QY	2048	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTG	2107
DB	3968	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTG	4027
QY	2108	GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACCTTGGTCCAGCATCT	2167
DB	4028	GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACCTTGGTCCAGCATCT	4087
QY	2168	CAGCACTTTCTTTCCAGTGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAACAAA	2227
DB	4088	CAGCACTTTCTTTCCAGTGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAACAAA	4147
QY	2228	GTGCCCCACTATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACA	2287
DB	4148	GTGCCCCACTATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACA	4207
QY	2288	GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAAGATCTCAGCTTATAGGACTGCC	2347
DB	4208	GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAAGATCTCAGCTTATAGGACTGCC	4267
QY	2348	ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA	2401
DB	4268	ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA	4321

RESULT 7
 AX538624
 LOCUS AX538624 11443 bp DNA linear
 DEFINITION Sequence 44 from Patent WO0229056.
 ACCESSION AX538624
 VERSION AX538624.1 GI:25271175
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE 1
 AUTHORS Chamberlain, J.S. and Harper, S.Q.
 TITLE Mini-dystrophin nucleic acid and peptide sequences
 JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
 THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
 FEATURES
 source Location/Qualifiers
 1..11443
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Synthetic"

Query Match 55.6%; Score 1335.6; DB 6; Length 11443;

Best Local Similarity 98.3%; Pred. No. 0;		Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;	
QY	1028	TGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACACTCATAGA	1087
Db	5702	TGGAGAGCATTCATAAAGGGTGAGTGAGCGAGAGGCTGCTTTTGGAAAGAACTCATAGA	5761
QY	1088	TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAA	1147
Db	5762	TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAA	5821
QY	1148	ACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG	1207
Db	5822	ACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG	5881
QY	1208	GGAGTAAAAGAGCTGATGAACAATGCCAAGACCTCCAAGGTGAAATTGAAGCTCACACA	1267
Db	5882	GGAGTAAAAGAGCTGATGAACAATGCCAAGACCTCCAAGGTGAAATTGAAGCTCACACA	5941
QY	1268	GATGTTTATCAACACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTTCC	1327
Db	5942	GATGTTTATCAACACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTTCC	6001
QY	1328	GATGATGCACTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT	1387
Db	6002	GATGATGCACTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT	6061
QY	1388	CGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGTTCTTGACCAGTGGAAAGCGT	1447
Db	6062	CGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGTTCTTGACCAGTGGAAAGCGT	6121
QY	1448	CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	1507
Db	6122	CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	6181
QY	1508	CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAAGAGCAGAACGATGTACATAGG	1567
Db	6182	CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAAGAGCAGAACGATGTACATAGG	6241
QY	1568	GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA	1627
Db	6242	GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA	6301
QY	1628	CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTACCAGGAGGCCAGA	1687
Db	6302	CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTACCAGGAGGCCAGA	6361
QY	1688	GAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG	1747
Db	6362	GAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG	6421
QY	1748	GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA	1807
Db	6422	GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA	6481
QY	1808	GATGAGACCTTTGAAAGACTCCAGGAATTCAGAGGCCACCGATGAGCTGGACCTCAAG	1867
Db	6482	GATGAGACCTTTGAAAGACTCCAGGAATTCAGAGGCCACCGATGAGCTGGACCTCAAG	6541
QY	1868	CTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC	1927
Db	6542	CTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC	6601
QY	1928	TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCAGAGGAAATTTGCGCCTCTGAAA	1987
Db	6602	TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCAGAGGAAATTTGCGCCTCTGAAA	6661
QY	1988	GAGAACGTGAGCCACGTCAAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC	2047
Db	6662	GAGAACGTGAGCCACGTCAAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC	6721
QY	2048	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTG	2107

6722 TCACCGTATAAACCTCAGCACCTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 6781
2108 GCCGTGAGGACCGAGTCAGGCACTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCT 2167
6782 GCCGTGAGGACCGAGTCAGGCACTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCT 6841
2168 CAGCACTTTCTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227
6842 CAGCACTTTCTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 6901
2228 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGGTGGGACCATCCCAAAATGACA 2287
6902 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGGTGGGACCATCCCAAAATGACA 6961
2288 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 2347
6962 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 7021
2348 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 2401
7022 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 7075

RESULT 8
AX538627 12057 bp DNA linear PAT 23-NOV-2002
LOCUS
DEFINITION
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
1
REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. 12057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 55.6%; Score 1335.6; DB 6; Length 12057;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
1028 TGGGCAACATCTCTGATGACAGAGACCGCTGGTCTTTTACAAGACACTCATAGA 1087
8216 TGGAGAAGCATTCATAAAAGGGTGAGTGAGGAGAGAGGCTGCTTTGGAAGAACTCATAGA 8275
1088 TTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAA 1147
8276 TTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAA 8335
1148 ACAACTGCCAATGTCTACAGGATGCTACCGTAAGGAAAGGCTCTAGAGACTCCCAAG 1207
8336 ACAACTGCCAATGTCTACAGGATGCTACCGTAAGGAAAGGCTCTAGAGACTCCCAAG 8395
1208 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 1267
8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 8455
1268 GATGTTTATCACAACTGGATGAAACACAGCCCAAAAAATCCCTGAGATCCCTGGAAGGTTCC 1327
8456 GATGTTTATCACAACTGGATGAAACACAGCCCAAAAAATCCCTGAGATCCCTGGAAGGTTCC 8515
1328 GATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 1387
8516 GATGATGCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 8575

QY 1388 CGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT 1447
Db 8576 CGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT 8635
QY 1448 CTGCACCTTTCTCTGACAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGC 1507
Db 8636 CTGCACCTTTCTCTGACAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGC 8695
QY 1508 CGGCAGGCACCTATTTGGAGGCGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGG 1567
Db 8696 CGGCAGGCACCTATTTGGAGGCGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGG 8755
QY 1568 GCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTA 1627
Db 8756 GCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTA 8815
QY 1628 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCACAGA 1687
Db 8816 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCACAGA 8875
QY 1688 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
Db 8876 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
QY 1748 GAGGTCAATACTGACTGGGAAAAATTTGAACCTTGCACCTCCGCTGACTGGCAGAGAAAAATA 1807
Db 8936 GAGGTCAATACTGACTGGGAAAAATTTGAACCTTGCACCTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1808 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGAGATGAGCTGGACCTCAAG 1867
Db 8996 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGAGATGAGCTGGACCTCAAG 9055
QY 1868 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGAC 1927
Db 9056 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGAC 9115
QY 1928 TCTCTCCAAGATCACTCTGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTGAAA 1987
Db 9116 TCTCTCCAAGATCACTCTGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTGAAA 9175
QY 1988 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTC 2047
Db 9176 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTC 9235
QY 2048 TCACCGTATAACCTCAGCACTCTTGGAAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG 2107
Db 9236 TCACCGTATAACCTCAGCACTCTTGGAAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG 9295
QY 2108 GCCGTGAGGACCGAGTCAGGCACTGCAATGAAGCCCAAGGACCTTTGGTCCAGCATCT 2167
Db 9296 GCCGTGAGGACCGAGTCAGGCACTGCAATGAAGCCCAAGGACCTTTGGTCCAGCATCT 9355
QY 2168 CAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227
Db 9356 CAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 9415
QY 2228 GTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACA 2287
Db 9416 GTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACA 9475
QY 2288 GAGCTCTACAGTCTTTTAGTGAACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 2347
Db 9476 GAGCTCTACAGTCTTTTAGTGAACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 9535
QY 2348 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCA 2401
Db 9536 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCA 9589

RESULT 9
HSDMDR
LOCUS

12446 bp mRNA linear PRI 12-SEP-1993

HSDMDR

DEFINITION Human mRNA for dystrophin.
ACCESSION X14298
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 12446)
REFERENCE Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.
AUTHORS Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR
COMMENT see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES
source Location/Qualifiers
1..12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosome, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99..11156
/note="unnamed protein product; dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="GOA:P11532"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVQKFTTKWNAQFSKFKQHLENLFSDL QDGRLLDLEGLTGQKLPEKSGSTRVHALNNVNKALRVLQNNVLDVNLIGSTDIVDG NHKLTILGIWNIIILHWQVKNVMKNIMAGLOPTNSEKILLSWVRQSRPNYPQVNVINF TSWSDGLALNALIHSRFPDLFDWNSVWCQSATORLEHAFNARIYQIGIEKLLHPDEDV DTPYDKKSILMYITSLFQVLPQQVSIEAIQEVEMLP RP PKVTKEHFQJLHQMHYSQ QITVSLAQGVERTSPKPRFSYAYTQAAVTTSDTPTRSPFPSOHLEAPEDKSFSGSL MESEVNDLRYQTALAEVLSWLLSABDTLQAQGEISNDVEVVKDQFTHGEGYMDLTAH QGRVGNILQGLSKLIGTGKLSDEDETEVQEQMNLNRSWECLRVASNEKQSNLHRVLM DLQNKLELNDLWLTKEERTRKMBEELPLGPDLEDLKRQVQKHVKVLEDELEQEVRTN SLTHMVFVVDES GDHATAALEQLKVLGDRWANTCRWTEDRWVLLQDILLKQWRLTE EQCLFSAWLSEKEDAVNKHHTGFKDQNEMLSSLSQKLAVLKADLEKKQSMGKLYSIK QDLSLILKNKSVTQKTEAWLNDNFARCDNLVQLEKSTAQISQAVITTPSLTQTVM ETVTVTVTREQILVKHAQEEPPPPQKKRQITVDSIEIRKRLDVIDITELHSWITRSEA VLQSPFAIFRKEGNFSDLKKNVNAIEREKAERFKLQDASRSGQALVEQMVNEGVA DSIKQASEQLNSRWIEFCQLISERLNLWLEYQNNI IAFYNQLQQLEQMTTAENWLKIQ PTPPSEPTAIKSQKICKDEVNRLSGLQPQIERLKIQSIALKEKGQGMFLDADFVAF TNHFKQVFSQVQAREKELQTFDLPMPRYQETMSAIRTWVQSQSETKLSIPQLSVTDY EIMEQRLGELQALQSSLOEQSGLYLSTTVKEMSKKAPSEISRKYQSEFEFEEIEGRWK KLSQLVEHCQKLEEQMNKRLKIQNHIQTLKWNMAEVDVFLKEEWPALGDS EILKKQL KQCRLLVSDIQTIQPSLNSVNEGQKIKNEAPEFASRLETETELKELNTQWDHMCQQVY ARKEALKGLEKTVSLQDLSBHEWMTQAEIEYLERDFFEYKTPDELQKAFPEEMKRAK EEAQQKEAKVLLTESVNSVIAQAPPVAQEAALKKELETLTITNYQWLCTRLNGKCKTLE EVWACWHELILLSYLEKANKWLNVEFFKLKTENIPGAEEISEVLDSLENLMRHSEDNP NQIRILAQTILTDGGVMDELINBELETFSNRWRLEHEEAVRRQKLEQSIQSAQETENS LHLIOESLTFIDKQLAAAYIADKVDAQMPQEAQKIQSDLTSHIEISLEEMKKHNQKGEA AQRVLSQIDVAQKQLQDVSMKFLRQKPFANFEQRIQESKMILDEVQWHLPALETKSVE QEVVQSLNHCNVNLYKSLSEVKSEVEMVIKTGRQIQVKQQTENPKELDERVATLKLHY NELGAKVTERKQQLKCLKLSRKMVKEMNVLTEWLAATDMELTKRSAVEGMPNSLDSE VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKETLVEDKLSLNSNWIAVTSRAE EWLNLLEYQKHMETFDQNVDHITKWIIQADTLDESEKKKPOQKEDVLKRLKAEIAND IRPKVDSNRDQANLMLANMANGHCKRLVEPQISELNRFPAAISHRIKTKGASIPLKELE QFNSDIQKLEPLEAEIQQGVNLKEEDFNKOMNEDNEGTVKELLQRGNLQQRITDER KSEIETIKQQLLTQKNALKDLRSQRRKKALEISHQWYQYKQADDLKCLDDIEKLL

ASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE IESKPAQFRRLNFAQIHTVREETMMVTEDMPEIISYVSTYLTFTIHTSVSQAALLEVEQ LNAAPDLCAKDFEDLFKQBEESLKNIKDSLQSSGRIDIHHSKTTAALQSATPVERVKL QEALSQDLDFQWEKVNMYKDRQGRFDRSVERWRFFHYDIKIFNQWLTEAEQFLRKTIQI PENWEHAKYKWLKELQDQIGQRTVTVTLNATGEEIIQOSSKTDSILQEKLGSLNL RWQEVCKQLSDRKKRLEEQNILSEFORDNLNFVLEEAADNIASIPLEPGKEQQLKE KLEQVKLLVEELPLRQGLKQLNETGGPVLVSAPISPEEQDKLENKLTQTNLQWIKVS RALPEKQGEIEAOIKDLGOLEKLELDEEQNLHLLWLSPIRNQLEINQPNQOEGPFD VKETEIAVQAKQPDVEEILSKGOHLYKEKPATQPVKRLKLELDSSEKAVNRLQLLELRA KQPD LAPGLTTIGASPTQTVTLVTPQVVTETAISKLEMPSSLMLEVPALADFNRAWT ELTDWLSLLDQVIKVSQVMVGDLEDINEMIKQKATMQDLEQRRPQLEELITAAQNLIK NKTSNQEARTIITDRIERIQNWDEVOEHLQNRQQLNEMLKQDSTOWLEAKEEAEQVL GQARAKLESWKGPYTVDAIQKKITETKQLAKDLRQWQNTVDVANDLALKLRDYSAD DTRKVMHITENINASWRSIHKRVSEREALEETHRLLLQFPDLDEKFLAWLTEAETTA NVLQDATRKERLLEDKSGVKELMKQWDLQGEIEAHTDVYHNLDENSQKILRSLEGSD DAVLLQRRLLDNMNFKWSELRKSLNIRSHLEASDQWKRLHLSLOELLVWLQKXDEL SRQAPIGGDFPAVQKQNDVHRAFKRELKTKPEVIMSTLETVRIFLTEQLEGLEKLYQ EPRELPEERAQNVTRLRKQAEEVTEWEKLNLSADWQKIDETLERLOELQEATD ELDLKLQRAEVIKGSWQPVGDLLIDSLQDLEKVKALRGEIAPLKENVSHVNDLARQL TTIGLISLPYNLSTLEDLNRWKLQVAVEDRVROLHEAHRDFGPASOHFLSTSVQGP WERAISPNKVPYIYNHETQTTCDWHFKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA LCIDLLSAACDALDQHNLKQNDQPMDILQIINCILTIYDRLEQEHNNLVNVPCLVD MCLNWLNVYDTRGTRIRVLSFKTGISLCKAHLEDKYRILFKQVASSTCFCDQRRRL GLLLHDSIQIPROLGEVASFSGSNIESPVSRCFOPANNKPEIEAALFLDMWRLEPQSM VMLPVLHRAVAAETAKEHQAKNICKEPIIGFRYRSLKHFNIDICQSCFFSGRVAKGH XMHYPMVEYCTPTTSGEDVDFPAKVLKNKERTKRVFAKHPRMGYLPVQTVLEGDNMET PVTLINFPFVDSAPASSPOLSHDDTHSRIEHYASRLAEMENSGSYLSDISPNESID DEHLIHYCQSLNQDSPLSQSPSPAQIILISLESEBERGELELILADLEENRNLAQEY DRLKQOHEHKGLSPLPSPPEMPTSPQSPRDAELIABAKLLRHQKGRLEARNQIILEDH NKQLESQJLHRLRQLLEQPOAEAKVNGTTVSSPSTLQSRDSSQPMLLRVRVGSQTSDSM GEEDLILSPPDQDTSTGLEEVNEQLNNSFPSSSRGRNTPGKPMREDTM"

ORIGIN

Query Match	55.6%;	Score 1335.6;	DB 9;	Length 12446;
Best Local Similarity	98.3%;	Pred. No. 0;		
Matches 1350;	Conservative	0;	Mismatches 24;	Indels 0; Gaps 0;
QY	1028	TGGGCAACATCTGTAGATGGACAGACACCGCTGGGTTCTTTTACAAGACACTCATAGA	1087	
Db	8106	TCGAGAAAGCATTCATAAAAGGGTGAGTGACCGGAGAGGCTGCTTTGGAAGAAACTCATAGA	8165	
QY	1088	TTACTGCAACAGTTCCTCCCTGGACCTGGAAGGTTCTTCTGCTGGCTTACAGAAGCTGAA	1147	
Db	8166	TTACTGCAACAGTTCCTCCCTGGACCTGGAAGGTTCTTCTGCTGGCTTACAGAAGCTGAA	8225	
QY	1148	ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCTCCTAGAAGACTCCAAG	1207	
Db	8226	ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGCTCCTAGAAGACTCCAAG	8285	
QY	1208	GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA	1267	
Db	8286	GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA	8345	
QY	1268	GATGTTTATCACAACTGGATGAAACAGCCATTTGGATAAACATGAACCTTCAAGTGGAGTTCC	1327	
Db	8346	GATGTTTATCACAACTGGATGAAACAGCCATTTGGATAAACATGAACCTTCAAGTGGAGTTCC	8405	
QY	1328	GATGATGCAGTCCTGTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCTT	1387	
Db	8406	GATGATGCAGTCCTGTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCTT	8465	
QY	1388	CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT	1447	
Db	8466	CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT	8525	
QY	1448	CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	1507	
Db	8526	CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	8585	
QY	1508	CGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGG	1567	
Db	8586	CGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGG	8645	

QY 1568 GCCTTCAAGAGGGAAATTGAAAACTAAAGAACTGTGTAATCATGAGTACTCTTTGAGACTGTGA 1627
|||||
Db 8646 GCCTTCAAGAGGGAAATTGAAAACTAAAGAACTGTGTAATCATGAGTACTCTTTGAGACTGTGA 8705
|||||
QY 1628 CGAATATTTCTGACAGAGCAGCCCTTTTGGAGGACTAGAGAACTCTTACCAGGAGCCGAGA 1687
|||||
Db 8706 CGAATATTTCTGACAGAGCAGCCCTTTTGGAGGACTAGAGAACTCTTACCAGGAGCCGAGA 8765
|||||
QY 1688 GAGCTGCCTCCTGAGGAGAGAGCCCGAGAATGTCTCACTCGGCTTCTACGAAAAGCAGGCTGAG 1747
|||||
Db 8766 GAGCTGCCTCCTGAGGAGAGAGCCCGAGAATGTCTCACTCGGCTTCTACGAAAAGCAGGCTGAG 8825
|||||
QY 1748 GAGGTCAATACTGAGTGGGAAAAAATTGAACTGCTCACTCCGCTGACTGGCAGAGAAAAATA 1807
|||||
Db 8826 GAGGTCAATACTGAGTGGGAAAAAATTGAACTGCTCACTCCGCTGACTGGCAGAGAAAAATA 8885
|||||
QY 1808 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
|||||
Db 8886 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 8945
|||||
QY 1868 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 1927
|||||
Db 8946 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 9005
|||||
QY 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGA 1987
|||||
Db 9006 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGA 9065
|||||
QY 1988 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 2047
|||||
Db 9066 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 9125
|||||
QY 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 2107
|||||
Db 9126 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 9185
|||||
QY 2108 GCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCCCACAGGCACTTTGGTCCAGCATCT 2167
|||||
Db 9186 GCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCCCACAGGCACTTTGGTCCAGCATCT 9245
|||||
QY 2168 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 2227
|||||
Db 9246 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 9305
|||||
QY 2228 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACA 2287
|||||
Db 9306 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACA 9365
|||||
QY 2288 GAGCTTACAGTCTTTAGCTGACCTGAATAATGTGATCTCAGCTTATAGGACTGCC 2347
|||||
Db 9366 GAGCTTACAGTCTTTAGCTGACCTGAATAATGTGATCTCAGCTTATAGGACTGCC 9425
|||||
QY 2348 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 2401
|||||
Db 9426 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 9479
|||||

RESULT 10
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source
Location/Qualifiers
1..13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"
/note="EMBL/GenBank Accession No. M18533"
ORIGIN
Query Match 55.6%; Score 1335.6; DB 6; Length 13957;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1028 TGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAAGACACTCATAGA 1087
|||||
Db 8216 TGGAGAAGCATTCATAAAAGGGTGAGTGAGCGAGAGGCTGCTTTTGGAGAAACTCATAGA 8275
|||||
QY 1088 TTACTGCAACAGTTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAA 1147
|||||
Db 8276 TTACTGCAACAGTTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAA 8335
|||||
QY 1148 ACACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1207
|||||
Db 8336 ACACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 8395
|||||
QY 1208 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
|||||
Db 8396 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 8455
|||||
QY 1268 GATGTTTATCACAACTCGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
|||||
Db 8456 GATGTTTATCACAACTCGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 8515
|||||
QY 1328 GATGATGCAGTCTCTCAACATTAGGTCCCATTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 1387
|||||
Db 8516 GATGATGCAGTCTCTCAACATTAGGTCCCATTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 8575
|||||
QY 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGAAGCGGT 1447
|||||
Db 8576 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGAAGCGGT 8635
|||||
QY 1448 CTGACCTTTCTCTGAGGAACTTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGC 1507
|||||
Db 8636 CTGACCTTTCTCTGAGGAACTTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGC 8695
|||||
QY 1508 CGGAGGCGACCTATTGGAGGCGACTTTTCCAGAGTTTCAAGAGCAGATGTACATAGG 1567
|||||
Db 8696 CGGAGGCGACCTATTGGAGGCGACTTTTCCAGAGTTTCAAGAGCAGATGTACATAGG 8755
|||||
QY 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAATCATGAGTACTCTTTGAGACTGTA 1627
|||||
Db 8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAATCATGAGTACTCTTTGAGACTGTA 8815
|||||
QY 1628 CGAATATTTCTGACAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACGAGGAGCCCAGA 1687
|||||
Db 8816 CGAATATTTCTGACAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACGAGGAGCCCAGA 8875
|||||
QY 1688 GAGCTGCCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG 1747
|||||
Db 8876 GAGCTGCCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG 8935
|||||
QY 1748 GAGGTCAATACTGAGTGGGAAAAAATTGAACTGCACTCCGCTGACTGGCAGAGAAAAATA 1807
|||||
Db 8936 GAGGTCAATACTGAGTGGGAAAAAATTGAACTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
|||||
QY 1808 GATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
|||||
Db 8996 GATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055
|||||
QY 1868 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGCTGGGCGATCTCCTCATTTGAC 1927
|||||
Db 9056 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGCTGGGCGATCTCCTCATTTGAC 9115
|||||
QY 1928 TCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCCTCTGAAA 1987
|||||

Db	9116	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCCTCTGAAA	9175
Qy	1988	GAGAACGTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCACTTCAGCTC	2047
Db	9176	GAGAACGTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCACTTCAGCTC	9235
Qy	2048	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG	2107
Db	9236	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG	9295
Qy	2108	GCCGTGAGGACCGAGTCAAGGCACTGCAATGAAGCCCAAGGACCTTGGTCCAGCATCT	2167
Db	9296	GCCGTGAGGACCGAGTCAAGGCACTGCAATGAAGCCCAAGGACCTTGGTCCAGCATCT	9355
Qy	2168	CAGCACTTTCTTCCACGCTGTCCAGGTCCTGAGGAGAGAGCCATCTCGCCAAACAAA	2227
Db	9356	CAGCACTTTCTTCCACGCTGTCCAGGTCCTGAGGAGAGAGCCATCTCGCCAAACAAA	9415
Qy	2228	GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGTGGGACCAATCCCAAAATGACA	2287
Db	9416	GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGTGGGACCAATCCCAAAATGACA	9475
Qy	2288	GAGCTCTACCACTTTTAGCTGACCTGAAATATGTCAGATTTCTCAGCTTATAGGACTGCC	2347
Db	9476	GAGCTCTACCACTTTTAGCTGACCTGAAATATGTCAGATTTCTCAGCTTATAGGACTGCC	9535
Qy	2348	ATGAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA	2401
Db	9536	ATGAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA	9589
RESULT 11			
AX538581			
LOCUS			
DEFINITION			
Sequence 1 from Patent WO0229056.			
ACCESSION			
AX538581			
VERSION			
AX538581.1 GI:25271086			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)			
FEATURES			
source			
1..13957			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1350; Conservative			
55.6%; Score 1335.6; DB 6; Length 13957;			
98.3%; Pred. No. 0;			
0; Mismatches			
24; Indels			
0; Gaps			
Qy	1028	TGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACACTCATAGA	1087
Db	8216	TGGAGAAGCATTCATAAAAGGTCAGTGAGCGAGAGGCTGCTTTGGAAGAACTCATAGA	8275
Qy	1088	TTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAA	1147
Db	8276	TTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAA	8335
Qy	1148	ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAG	1207
Db	8336	ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAG	8395
Qy	1208	GGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTGAGCTCACACA	1267

Db	8396	GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA	8455
Qy	1268	GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCC	1327
Db	8456	GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCC	8515
Qy	1328	GATGATGAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT	1387
Db	8516	GATGATGAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT	8575
Qy	1388	CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGAAGCGT	1447
Db	8576	CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGAAGCGT	8635
Qy	1448	CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	1507
Db	8636	CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	8695
Qy	1508	CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG	1567
Db	8696	CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG	8755
Qy	1568	GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA	1627
Db	8756	GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA	8815
Qy	1628	CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAGGAGCCCAGA	1687
Db	8816	CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAGGAGCCCAGA	8875
Qy	1688	GAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG	1747
Db	8876	GAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG	8935
Qy	1748	GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA	1807
Db	8936	GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA	8995
Qy	1808	GATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	1867
Db	8996	GATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	9055
Qy	1868	CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC	1927
Db	9056	CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC	9115
Qy	1928	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA	1987
Db	9116	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA	9175
Qy	1988	GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCACTTCAGCTC	2047
Db	9176	GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCACTTCAGCTC	9235
Qy	2048	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG	2107
Db	9236	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG	9295
Qy	2108	GCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACTTTGGTCCAGCATCT	2167
Db	9296	GCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACTTTGGTCCAGCATCT	9355
Qy	2168	CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA	2227
Db	9356	CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA	9415
Qy	2228	GTGCCCTACTATATCAACCAAGACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACA	2287
Db	9416	GTGCCCTACTATATCAACCAAGACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACA	9475
Qy	2288	GAGCTCTACCAAGTCTTTAGCTGACCTGAAATAATGTCAGATTTCTCAGCTTATAGGACTGCC	2347
Db	9476	GAGCTCTACCAAGTCTTTAGCTGACCTGAAATAATGTCAGATTTCTCAGCTTATAGGACTGCC	9535

2348 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCGTGCA 2401
|||||
9536 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCGTGCA 9589
|||||

RESULT 12
HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
LOCUS Homo sapiens dystrophin (DMD) mRNA, complete cds.
DEFINITION M18533 M17154 M18026 M20250
ACCESSION M18533.1 GI:181856
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1699)
Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.
Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
Cell 50 (3), 509-517 (1987)
87273512
3607877
2 (bases 1678 to 3830)
Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
Conservation of the Duchenne muscular dystrophy gene in mice and humans
Science 238 (4825), 347-350 (1987)
88018015
3659917
3 (bases 1 to 13957)
Koenig,M., Monaco,A.P. and Kunkel,L.M.
The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein
Cell 53 (2), 219-226 (1988)
88194521
3282674
On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.
Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEDCYEREDYQKKTFTKWVNAQFSKFGKHLENLFSDL QGRRLLDLLEGLTGQKLPKEKGSSTRVHALNNVKNALRVLQNNVDLVNIGSTDIVDG NHKLTGLIWNIIHLHWQVKNVMKNIMAGLQQTNSEKILLSWYRQSTRNYPQVNVINPT TSWSDGLALNALIHSHRPDLFDWNSVVCQOSATQRLAHAFNARYQLGIEKLLDPEDV DTTYPDKSILMYITSLFQVLPQQVSIETAEVEMLPRPKVTKEEHFLHHQMHYSQ QITVSLAQGYERTSPKPRFKSYATQAYVTSDPTRSPFSQHLEAPEDSKSFGSSL MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLTAH QGRVGNILQGLCTGKLSDEDETEVQEQMNLNLSRWECILRVASMEKQSNLHRVLM DLQNQKLELNDWLTKEERTRKMEEEPGLGPDLEDLKRQVQQHKVQLQEDLEQEQVRVN

FEATURES
source

gene
CDS

SLTHMVVVDESSGDHATAALEEQKVLGDRWANICRWTEDEWVLLQDILLKWQRLTE
BQCLPSAWLSEKEDAVNKIHTTGFKDONEMLSLQKLAVLKADLEKKKQSMGKLYSLK
QDILLSTLKNKSVTKTEAWLDNFARCDWNLVQKLEKSTAQISQAVTTTQPSLTQTITVM
ETVTVTTREQIILVHAQEELPPPPQKKRQITVDSEIRKRLDVIDTELHSWITRSEA
VLOSPEFAIFRKEGNSDLKEKVNAIEREKAERFKRLQDASRAQALVEQVMNEGUNA
DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
PTTPSEPTAIKSQKICKDEVNLSGLQPOIERLKIQSIALKEKGQGPMLDADFVAF
TNHFQVFSVDQAREKELQTIPTDLPFPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
EIMEQRLGELQALQSSIQEQSGLYLSTVTKEMSKAPSEISRKYQSEFEEIEGRWK
KLSSQLVEHCQLEQMNKLRKIQNHIIQTLKXMAEVDVFLKEEWPALGDSIILKKQL
KOCRLIVSDIQTIQPSLNSVNEGQKIKNEAPEFASRLTELKELNTQWDMHMCQQVY
ARKEALKGLEKTVSLQKDLSEMHWMTOAEEYLERDFEYKTPDELQKAVEEMKRAK
EEAQOKEAKVKLLTESVNSVIAQAPFVACALKKELETLTTNYQWLCRLNGCKTLE
EVMACWHELLSYLEKANKWLNVEFEKLTETENIPGGAEEISEVLDLENLMRHSNDP
NOIRILAOITLTDGGVMDLINELELETFSNRWELHEEAVRRQKLLFQSIQSAQETKS
LHLIQESLTFIDKQLAAYIADKVDAQMPQEAQKIOSDLSHSIEISLEMKHNGQKEA
AGRVLSQIDVAOKLQDVSMKFRLLFOKPAFELRLQESKMILDEVKXHLPALETKSVE
QEVVQSLNHCVNLYKSLSEVKSEVMVKTGRQIVQKKQTEPNKPELDERVETALKLHY
NELGAKVTERKQOQLEKCLKLSRKMRKEMVNLTEWLAATDMELTKRSAVEGMPNSLDSE
VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKKTLEVEDKLSLNSNWIATVTSRAE
EWLNLLEYQKHMETFDONVDHITKWIIOADTLLDESEKKPQKEDVLKRLKAEKND
IRPKVDSTRDQAANLMANRGDHCRLKVEPQISELNRFAAISHRKICGKASIPKLELE
QFNSDIQKLEPLEAEIQQGVNLKEEDFNKMDNEDNEGTVKELLORGDNLQQRITDER
KREEIKIKQQLLTQKHNAKDLRSQRRKKALEISHQWYQYKQADDLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKEELNAVRRQAELSGEAGMAVAVEPTQIQLSKRWRE
IESKFAQFRRLNFAQHTVREETMMVTMDPLEISYVPSTLTETITHVSQALLEVEQ
LNAAPDLCAKDFEDLFKQEESLKNIKDSLQOSSGRIDIHSKTAALQSATPVERVKL
QEALSQLDFQWEKVMKYDRQGRFDRSVEKWRFRHYDIKIFNQLWTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGQROTIVRTLNTATGEEIIQOSSKTDASILQEKLSNL
RWQEVCKQLSDRKRELEEKNILSEFQDLNFEVLWEEADNIIASIPLEPKEQQLKE
KLEQVKLLVEELPLRQGIKQLNETGGPVLVSAPISPEEQDKLENKLTQNLQWIKVS
RALPEKQGEIEAQIKDLGLEKLEDELEQNLHLLWLSPIRNQLEIYNQPNQEGPFD
VQETIAVQAKQPDVEEILSKGQHLXKEKPAQPVKRLKEDLSSEWKAVNRLLOELRA
KOPDLAPGLTTIGASPTQTVLTPVPTVTKETAISKLEMPSSLMEVPEALADFNRAWT
ELTDWLSLLDQVKSQRMVGDLEDINEMIIOKATMQDLEQRRPQLEELITAAQNKL
NXTSNQEAARTIITDRIERIQNQWDEVQEHQNRQQLNEMLKDSTOWLEAKBEAEQVL
GQARAKLESWKEGPTYVDAIQKKITETKQAKDLRQWQTNVDVANDLALKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSERAALBEETHRLLLQQFPDLDEKFLAWLTEAETTA
NVLQDATRKERLLEDSSKVKELMKQWQDLOGEIEAHTDVYHNLNDSQKILRSLEGSD
DAVLLQRRLLDNMFKWSLIRKSLNIRSHLPASSDQWKRLHLSQELLVWLQLKDDDEL
SROAPIGGDFPAVQKQNDVHRAFKRELKTKPEVIMSTLETVRIFLTQPLEGLEKLYQ
EPRELPEERAQONVTRLIRKQAEVENTEWELNLSADWQKIDETLERIQELQEAATD
ELDKLRQAEVKGWQVGDLLIDSLQHLKVKALRGWGEIAPLKNVSHVNDLARQL
TTLIGIQLSPYNLSTLEDLNRWKLLOQAVEDVRQLHEAHRDFGPASQHFSLSTSVQGP
WERAISPNKVPYYINHETQTTCDWHPKMTELYQSLADLNNRVFSAYTAMKLRRLQKA
LCJDLLSLSAACDALDOHNLKQNDQMDILOINCLTIYDRLQEQHNNLVNVPLCVD
MCLNWLNNVYDTRGTRIRVLSFKTGIISLCAHLEDKRYLFKQVASTGFCQDQRRRL
GULLHDSIQIPQLGCVASFGGSNIEPSVRSQCFANNKPEIAEALFLDMWRLPEQSM
VWLPVLRHVAAAEATAKQAKCNICECPIIGFIRYSLKHFNIDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTRKYFAKHPRMGYLPVQTVLEGDNMET
PVTLLNFWPVDSPASSQLSHDDTHSRIEHYASRLAEMENSNYSINDSISPNSID
DEHLIIQHYCQSLNQDSPSQSPSPAQIILISEEERGELERILADLEENRNQLQAEY
DRLKQOHEHKGSLPLSPSPMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQOILEDH
NKQLESQHLRLQLLEQPPQAEAKVNGTTVSSPSTLSQRSOSSQPMLLRVLVVGSTSDSM
GEEDLLSPQDQDTSTGLBEVMEQLNNSFPSSRGNTPGKPMREDTM"

ORIGIN

Query Match	55.6%;	Score 1335.6;	DB 9;	Length 13957;
Best Local Similarity	98.3%;	Pred.No. 0;		
Matches 1350;	Conservative 0;	Mismatches 24;	Indels 0;	Gaps 0;
Qy	1028	TGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACACTCATAGA	1087	
Db	8216	TGGAGAAGCATTCATAAAAGGTTGAGTGAGCGAGCGCTGCTTTTGGAGAAACTCATAGA	8275	
Qy	1088	TTACTGCAACAGTTCCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGAGCTGAA	1147	
Db	8276	TTACTGCAACAGTTCCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGAGCTGAA	8335	
Qy	1148	ACAACTGCCAAATGTCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGAGCTCCAAG	1207	
Db	8336	ACAACTGCCAAATGTCTACAGGATGCTACCCGTAAGGAAGGCTCTCTAGAGAGCTCCAAG	8395	

QY 1208 GGAGTAAAGAGCTGATGAAA CAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
Db 8396 GGAGTAAAGAGCTGATGAAA CAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 8455
QY 1268 GATGTTTATCACAACCTGGATGAAAA CAGCAAAAAATCCTGAGATCCCTTGGAAAGTTCC 1327
Db 8456 GATGTTTATCACAACCTGGATGAAAA CAGCAAAAAATCCTGAGATCCCTTGGAAAGTTCC 8515
QY 1328 GATGATGCAGTCTGTACAAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTT 1387
Db 8516 GATGATGCAGTCTGTACAAAAGACGTTTGGATAAACATGAACCTCAAGTGGAGTGAACCTT 8575
QY 1388 CGGAAAAAGTCTCTCAACATTAGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 1447
Db 8576 CGGAAAAAGTCTCTCAACATTAGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 8635
QY 1448 CTGCACCTTTCTCTCAGGAACTTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGC 1507
Db 8636 CTGCACCTTTCTCTCAGGAACTTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGC 8695
QY 1508 CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGCAAGTTCAGAGCAGACCATGTA 1567
Db 8696 CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGCAAGTTCAGAGCAGACCATGTA 8755
QY 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGASTACTCTTGAGACTGTA 1627
Db 8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGASTACTCTTGAGACTGTA 8815
QY 1628 CGAATATTTCTGACAGAGAGAGCCAGAACTTTGGAAAGACTAGAGAACTCTACCAAGGAGCCAGA 1687
Db 8816 CGAATATTTCTGACAGAGAGAGCCAGAACTTTGGAAAGACTAGAGAACTCTACCAAGGAGCCAGA 8875
QY 1688 GAGCTGCCTCCTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
Db 8876 GAGCTGCCTCCTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
QY 1748 GAGGTCAATACTGAGTGGGAAAAATTGAACTTGCACCTCCGCTGACTGGCAGAGAAAAATA 1807
Db 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACTTGCACCTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1808 GATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
Db 8996 GATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055
QY 1868 CTGGCCCAAGCTGAGTGATCAAGGGATCCTTGGCAGCCCGTGGSGGATCTCCTCATTTGAC 1927
Db 9056 CTGGCCCAAGCTGAGTGATCAAGGGATCCTTGGCAGCCCGTGGSGGATCTCCTCATTTGAC 9115
QY 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTGCGCCTCTGAAA 1987
Db 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTGCGCCTCTGAAA 9175
QY 1988 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC 2047
Db 9176 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC 9235
QY 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG 2107
Db 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG 9295
QY 2108 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 2167
Db 9296 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 9355
QY 2168 CAGCACTTTCTTTCCAGTCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227
Db 9356 CAGCACTTTCTTTCCAGTCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 9415
QY 2228 GTGCCCTACTATATCAACACAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACA 2287
Db 9416 GTGCCCTACTATATCAACACAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACA 9475
QY 2288 GAGCTCTACCAGTCTTTAGCTGACCTGAAATATGTGAGATTCTCAGTATATAGGACTGCC 2347

Db 9476 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCC 9535
QY 2348 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTCA 2401
Db 9536 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTCA 9589
RESULT 13
AR220819 LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source 1..13977
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 55.6%; Score 1335.6; DB 6; Length 13977;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1028 TGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
Db 8216 TGGAGAAGCATTCATAAAAGGGTGAGTGAGCGAGAGGCTGCTTTTGGAAAGAACTCATAGA 8275
QY 1088 TTACTGCAACAGTTCCCGCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAA 1147
Db 8276 TTACTGCAACAGTTCCCGCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAA 8335
QY 1148 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAG 1207
Db 8336 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAG 8395
QY 1208 GGAGTAAAGAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
Db 8396 GGAGTAAAGAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 8455
QY 1268 GATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
Db 8456 GATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 8515
QY 1328 GATGATGCAGTCTCTGTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCT 1387
Db 8516 GATGATGCAGTCTCTGTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCT 8575
QY 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 1447
Db 8576 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 8635
QY 1448 CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
Db 8636 CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
QY 1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCCAGCAAGTTCAGAGCAGACCATGTA 1567
Db 8696 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCCAGCAAGTTCAGAGCAGACCATGTA 8755
QY 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGACTCTTTGAGACTGTA 1627
Db 8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAAATCATGACTCTTTGAGACTGTA 8815
QY 1628 CGAATATTTCTGACAGAGAGAGCTTTTGGAAAGACTAGAGAACTCTACCAAGGAGCCAGA 1687

```
||||| 8816 CGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGAGCCCAGA 8875
||||| 1688 GAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
||||| 8876 GAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
||||| 1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCGAGAAAAATA 1807
||||| 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCGAGAAAAATA 8995
||||| 1808 GATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGSCCAGCGATGAGCTGGACCTCAAG 1867
||||| 8996 GATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGSCCAGCGATGAGCTGGACCTCAAG 9055
||||| 1868 CTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGAC 1927
||||| 9056 CTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGAC 9115
||||| 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAA 1987
||||| 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAA 9175
||||| 1988 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTC 2047
||||| 9176 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTC 9235
||||| 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCATGGAAGCTTCTGCAGGTG 2107
||||| 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCATGGAAGCTTCTGCAGGTG 9295
||||| 2108 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGACTTTGTGTCAGCATCT 2167
||||| 9296 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGACTTTGTGTCAGCATCT 9355
||||| 2168 CAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCCCAAAACAAA 2227
||||| 9356 CAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCCCAAAACAAA 9415
||||| 2228 GTGCCCTACTATATCAACCACGAGACTCAAAACACTTGCTGGGACCATCCCAAAATGACA 2287
||||| 9416 GTGCCCTACTATATCAACCACGAGACTCAAAACACTTGCTGGGACCATCCCAAAATGACA 9475
||||| 2288 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 2347
||||| 9476 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 9535
||||| 2348 ATGAAACTCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCA 2401
||||| 9536 ATGAAACTCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCA 9589
```

```
RESULT 14
AF070485 AF070485 13887 bp mRNA linear MAM 09-DEC-1998
LOCUS Canis familiaris dystrophin mRNA, complete cds.
DEFINITION AF070485
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 13887)
AUTHORS Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES
location/Qualifiers
source 1..13887
organism="Canis familiaris"
mol_type="mRNA"
strain="breed Golden retriever"
```

```
/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="dystrophin"
/protein_id="AACB3646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEDCYEREDVQKFTTKWVNAQSKFGKOHENLFSDL
QDGRLLDLLEGLTGQKLPEKGSSTRVHALNNVKALRVLQKNVDLVNIGSTDIVDG
NHKLTGLIWNIIILHWQVKVMKNI MAGLQCTNSEKILLSWVRQSTRNYPQVNVNFT
TSWSDALNALNIHSHRPDLFDWNSVVCQSQATQREHAFNIAKYQLGIEKLLDPEDV
ATTYPDKSILMYITSLFOVLQQQVIEAIQEVEMLPRPSQVTRREEHFIHQHMHSQ
QITVSLAQGYERAPSPFPRFKSYAYTQAAVTTSDPTRSPLPSQHLETPEDKSFGRS
LTETEANLDSYOTALLEVLSWLLSAEDALQAQGEISNDVEEVKEQFTHHEGYMMDLTS
HQRVGNVLQSGQLIGTKLSEDETEVEQNMNLSRWECLRVASMEKQSNLHKVL
MDLQNOQLKELNDWLTKEERTRKMEKEPLGPDIEDLKROVQQHKVLOEDLEQEQVRV
NSLTHMVVVVDESSGDHATAALEQLKVLGDRWANI CRWTEDRWVLLQDILLKWRFT
EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVLSNLQKLAVLKTDLEKKKQMDKLCSL
NQDLISALKNTVVAHKMEAWLNDNFAQRWDNLVQKLEKSSAQISQAVTTQPSLTQTTV
METVTMTTREHILVKHAQEBELPPPPQKKRQIIVDSEIRKRLDVIDITELHSWITRSE
AVLQSPFAIYRKEGNSDLKEKVNAIEREKAERPKLQDASRAQALVEQMVNNEGVN
ADSIQAASEQLNSRWIEFQQLSERLNWLEYQNNIIITFYNLQQLQEQMTTAEWNLKT
OPTTTSPTAIKSQKICKDEINRLSALQPIERLKIQSIALKEKGQGNPLDADFVA
FTNHENQVFADVAQAREKELQTIIFDSLPPMRYQETMSITLTWIIQOSETKLSIPQVTVE
YDIMEORLGELOALQSSLOEQOGLNYLSTTVKMSKKAPLSDISRKYQSEFEIEGR
WKLSSQLVEHCQKLEEQMAKLKIQNHIKTLKWI TEVDVFLKEEWPAIGDSEILKR
QLKOCRLVNDIOTIQPSLNSVNEGAKMKNEAPEFAGRELTELRELNTQWDMCRQ
VYARKEALKGGLDKTVLSQKDLSEMHWMTOAESEYLERDFFEYKTPDELQTAVEEMKR
AKEEAQOKEAKVKLLTESVNSVIAQAPPAQAEALKEKLDLTLTNYOQLTRINGCKRT
LEEVAWCHHELLSYLEKANKWLSEVEVKLKT TENISGGAEEIAEVLDSLENLMQHS
NPNOIRILAOQLTLDGGVMDELINEELETFNSRWELHHEEAVRRQKLEQSIQSAQIE
KSLHLIQESLSSIDKQLAAYIADKVDAQAAMPQEAQIKQSDLTSHESILEEMKKHNQK
ETAQRLSQIDVAQKLOQVSMKFRLLFOKPAFQRLQESKMILDEVKMHLPALETKS
VEQEVVQSLNHCNLYKLSSEVKSEVMVITGRQIVQKKQ TENPKELDERVTALKL
HYNEIGAKVTERKQOQLEKCLKLSRKMKEVNALTEWLAATDMELTKRSAVEGMPSNLD
SEVAGKATQKEIEKQKVHLKSVTEVGEALKTVLGKKEMLVEDKLSLNSNWIARTSR
AEEWNLILLEQKHMETFDQNVYITNWI IQADALLDESEKKKPPQKILKRIKAE
NDIRPKVDSTRDQAANLMANRGDHCRCVPEPKI SELNHRFAAISHRIKTGKASIPLKE
LEQFNSDIQKLLPLEAEIIQQGVNLKEEDFNKMSDNEGTVKELLQRGDNLQQRITD
ERKRBEIKIKQQLLQTKHNALKDLRSQRRKKALEISHQVYQYKRAQADDLLKCLDIEK
KLASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSEDDGAAMAVEPTQIQLSKRW
REIESKFAQFRRNLNFAQIHTVHEESVAMTDEMPLEISYVPSTYLTETITVHSQALSEV
EELNAPDLCAQDFEDLFKEEESLKNIKDSLQIISGRIDIHNKKTAAHLSATPAERA
KLQEALSRLLDFOWERVNNMYKDRQGRFDRSVEKRRFHYDMKILNQWLTEAESQFLKKT
VSRNLPEKQEEIEAHVVDLQLEEQNLHLLWLSPIRNQLEIYNQPNOTGPPDIKEIE
VAVQAQDPDVEGILSKGOHLYKEKPATQPAKRLLEDLSSDKWVVTQLLOELRAKQPGP
APGLTTRAPPSQTVTLVTPAVTKETAISKLEMPSSLLLEV PALADFNRAWTETLDW
LSLLDRVIKSQRMVVGDLIEDINEMI IKQATLOLEQRRRQLEELITAAQNLKNKTSN
QEARTIITDRIERIQSQWDEVQEHQNRRLQLTEMLKDSQTOWLEAKEEABQVLGQARA
KLESWEAPYTVDAIQKKITETKQALQKRLQWQINVDVANDLALKLLRDYSADDTRKV
HMITENINASWASIHKRLSREAALEETHRLLQOFPDLDEKFLAWL TEAETTANVLQD
ATHKERLLEDKSGVRELMKQWDLQGEIEAHTDIYHNLDENGQKVLSLEGSDDAALL
QRRLDNMNFKWSELRKKSILNIRSHLEASSDQWKLHLSLQELLVWLQKDDDEL SRQAP
IGGDFPAVQKQNDVHRAFRKELTKTEPVMSTLETVRIFLTEQPLEGLEKLYQEPREL
PPEERAQNVTRLRLRQAEVNTQWEKLVNHSADQWKIDEALERLQELQEAATDEL DLK
LRQAEVIKGSWQPVGDLILDSLDHLEKVKALRGEITPLKENVSYNLDARQLTTLGI
QLSPYNLTLEDLNTRWKLLQVAIEDRIRQLHEAHRDFGPASQHFSTSVQGPWERAI
LSLSAACDALDOHNLKONDQPMDILQVINCLTIIYDRLEQEHNNLVNVLVCDMCLNL
LLNVYDTRGTRIRVLSFKTGIISLCKRAHLEDKRYLFQKVASSTGFCQDRLGLLH
DSIQIPRQLGEVASFGGSNIEPSVRS CFQFANNKPEIEAALFLDWMRLPEPQSMVWLPV
LHRVAAAEATAKHOAKNICKECPIIGFRYRSLSKHFNYDICQSCFFSFGVAKGHMHP
MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKPRMGYLPVQTVLEGDNMETPVTLI
NFWPVDAPASSQLSHDDTHSRIEHYASRLKHMNSNGSLYNDISPNESIDDEHLL
IQHYWRSNLQESPLSQRPSPAQILISLESEERGELERILADLEGRRNINQAEYDRLKQ
QHEHGLSPLSPPEMMPTSPQSPRDAELIAEAKRLHQKRLRLEARMQILEDHKNQLE
SQLHRLRQLLEQPAQAEAKVNGTTVSSPSTSLQSDSSQPMLLLRVVGQSQTSSEMGEDL
LSPQDQSTSTGLEEVMQNLNHSFPSSSRGRNTPGKPMREDTM"
```

```
3'UTR
11333..>13887
11789..11828
repeat_region
```


Search completed: April 5, 2004, 02:39:52
Job time : 6272.59 secs

Db	8447	GATACTATCACAATCTTGATGAAATGGCCAAAAATCCTGAGATCCCTGGAAGGTTGCG	8506
2Y	1328	GATGATGAGTCTCTGTTACAAAAGACGTTTGGATAAACATCAAACTTCAAGTGGAGTGAAC	1387
Db	8507	GATGAAGCAGCCCTGTTACAAAAGACGTTTGGATAAACATCAAACTTCAAGTGGAGTGAAC	8566
2Y	1388	CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT	1447
Db	8567	CAGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCAAGTTCTGACCAAGTGAAGCGT	8626
2Y	1448	CTGCACCTTTCTCTCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	1507
Db	8627	TTGCATCTTTCTCTCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGAATGAGC	8686
2Y	1508	CGGCAGGCACCTATTTGGAGGCGACTTTCCAGCAGTTTCCAGAAACGATGATACATAGG	1567
Db	8687	CGTCAGGCACCCATCGGTGGTGTATTTCCAGCAGTTTCCAGAAACGATGATACATAGG	8746
2Y	1568	GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA	1627
Db	8747	GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTG	8806
2Y	1628	CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCCGA	1687
Db	8807	AGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCCGA	8866
2Y	1688	GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG	1747
Db	8867	GAACTGCCTCCTGAAGAAAGAGCTCAGAAATGTCACTCGGCTCCTACGAAAGCAGGCTGAA	8926
2Y	1748	GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA	1807
Db	8927	GAGGTCAACGCTGAATGGGACAAATTGAACCTGCGTCTGAGTATTGGCAGAGAAAAATA	8986
2Y	1808	GATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	1867
Db	8987	GATGAAGCTCTTGAAAGACTCCAGGAACTTCAAGGAGCTGCCGATGAATGGACCTCAAG	9046
2Y	1868	CTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGAC	1927
Db	9047	TTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCAGTGGGGGATCTCCTCATTGAC	9106
2Y	1928	TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAA	1987
Db	9107	TCTCTGCAAGATCACCTTGAAAAAGTCAAGGCACTTCGGGGAGAAATTGCACCTCTTAAA	9166
2Y	1988	GAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTC	2047
Db	9167	GAGAAATGTCAATCGTGTCAATGACCTTGCACATCAGCTGACCCACACTGGGCATTTCAGCTC	9226
2Y	2048	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG	2107
Db	9227	TCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATGGAGGCTTCTACAGGTG	9286
2Y	2108	GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT	2167
Db	9287	GCTGTGGAGGACCGTGTACAGCAGCTGCATGAAGCCACAGGGACTTTGGTCTGATCC	9346
2Y	2168	CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA	2227
Db	9347	CAGCACTTCTTTCCACTTCAGTTCAGGTCCCTGGGAGAGAGCCATCTCACCAACAAA	9406
2Y	2228	GTGCCCTACTATATCAACCCAGAGACTCAAAACACTTGTCTGGGAGAGAGCCATCCCAAAATGACA	2287
Db	9407	GTGCCCTACTATATCAACCCAGAGACCCAAACACTTGTCTGGGAGAGAGCCATCCCAAAATGACA	9466
2Y	2288	GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTACAGATTTCTCAGCTTATAGGACTGCC	2347
Db	9467	GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTACAGTTCCTCCGGTATAGGACTGCC	9526
2Y	2348	ATGAAACTCCGAAGACTGCAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCTGTCA	2401
Db	9527	ATGAAAGCTCAGAAGGCTCCAGAAGGCCCTTTTGGCTTGGATCTCTTGAGCCTGTCA	9580

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 11:55:32 ; Search time 619.552 Seconds
(without alignments)
16463.377 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 acatgcattcaacatcgcca.....tggatctcttgagccctgtca 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2401	100.0	3858	6	AAD37237	Aad37237 Human dys
2	2401	100.0	4825	6	AAD37257	Aad37257 Adeno-ass
3	2401	100.0	4848	6	AAD37263	Aad37263 Adeno-ass
4	2401	100.0	5060	6	AAD37264	Aad37264 Adeno-ass
5	2250	93.7	3999	6	AAD37234	Aad37234 Human dys
6	2250	93.7	4966	6	AAD37256	Aad37256 Adeno-ass
7	2250	93.7	4990	6	AAD37262	Aad37262 Adeno-ass
8	2067	86.1	4182	6	AAD37230	Aad37230 Human dys
9	2067	86.1	5149	6	AAD37255	Aad37255 Adeno-ass
10	1839.4	76.6	5462	6	ABK81999	Abk81999 DNA encod
11	1737	72.3	3531	6	AAD37238	Aad37238 Human dys
12	1737	72.3	4498	6	AAD37258	Aad37258 Adeno-ass
13	1725	71.8	5339	6	ABK81998	Abk81998 DNA encod
14	1695	70.6	3510	6	AAD37240	Aad37240 Human dys
15	1695	70.6	4476	6	AAD37259	Aad37259 Adeno-ass
16	1419	59.1	5417	6	ABK81997	Abk81997 DNA encod
17	1335.6	55.6	5952	5	AAD06794	Aad06794 Human dys
18	1335.6	55.6	8689	6	ABK82000	Abk82000 DNA encod
19	1335.6	55.6	11058	6	AAD37229	Aad37229 Human dys
20	1335.6	55.6	11241	6	ABK82005	Abk82005 cDNA encod
21	1335.6	55.6	11443	6	ABK82002	Abk82002 DNA encod
22	1335.6	55.6	12923	1	AAN90338	Aan90338 Sequence
23	1335.6	55.6	13957	6	ABK81959	Abk81959 cDNA encod

24	1335.6	55.6	13957	6	ABT10904	Abt10904 Human bre
25	1335.6	55.6	13957	6	ABN95786	Abn95786 Gene #228
26	1335.6	55.6	13957	6	ABS69900	Abs69900 Human dys
27	1335.6	55.6	13977	6	ABS70403	Abs70403 Human bon
28	1323	55.1	2169	6	AAD37232	Aad37232 Human dys
29	1263	52.6	4414	6	AAD37260	Aad37260 Adeno-ass
30	1251	52.1	3446	6	AAD37242	Aad37242 Human dys
31	1143.6	47.6	13815	2	AAV18885	Aav18885 Mus muscu
32	1143.6	47.6	13815	6	ABK81960	Abk81960 cDNA enco
33	1143.6	47.6	13815	6	ABI99799	Abi99799 Mouse isc
34	1143.6	47.6	19307	2	AAT27558	Aat27558 Shuttle v
35	1093.8	45.6	4402	3	Aaz48568	Aaz48568 A rod sho
36	1092.2	45.5	1991	6	AAD37231	Aad37231 Human dys
37	1077	44.9	1667	6	AAD37235	Aad37235 Human dys
38	1021.2	42.5	4402	3	AAZ48567	Aaz48567 A rod sho
39	1018.8	42.4	4075	3	AAZ48569	Aaz48569 A rod sho
40	976	40.6	1821	6	AAD37241	Aad37241 Human dys
41	750	31.2	1340	6	AAD37239	Aad37239 Human dys
42	709.2	29.5	3747	3	AAZ48566	Aaz48566 A rod sho
43	633.2	26.4	3275	1	AAN97129	Aan97129 Partial s
44	589	24.5	1434	6	AAD37243	Aad37243 Human dys
45	498.6	20.8	10705	7	ABT41896	Abt41896 Toxicity

ALIGNMENTS

RESULT 1
AAD37237
ID AAD37237 standard; DNA; 3858 BP.

XX AAD37237;

XX 21-AUG-2002 (first entry)

DE Human dystrophin minigene delta3849.

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

OS Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.

XX 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.

XX Example 1; Page 48-49; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC	dystrophy (BMD) in a mammalian subject. The present sequence is human
CC	dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC	hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC	and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX	
SQ	Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;
	Query Match 100.0%; Score 2401; DB 6; Length 3858;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ACATGCAATCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
DB	
600	ACATGCAATCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 659
QY	61 TGTGATACCACTATCCAGATAAGAGTCCATCTTAATGTACATCATCACTCTTCCA 120
DB	
660	TGTGATACCACTATCCAGATAAGAGTCCATCTTAATGTACATCATCACTCTTCCA 719
QY	121 AGTTTTGCCTCAACAAGTGAGCAATTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC 180
DB	
720	AGTTTTGCCTCAACAAGTGAGCAATTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC 779
QY	181 ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 240
DB	
780	ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 839
QY	241 GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAAGCCTCGATTCAA 300
DB	
840	GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 899
QY	301 GAGCTATGCCCTACACACAGCGTGTCTATGTACACCCTCTGACCCCTACACGGAGCCCAATT 360
DB	
900	GAGCTATGCCCTACACACAGCGTGTCTATGTACACCCTCTGACCCCTACACGGAGCCCAATT 959
QY	361 TCCTTCACAGCAATTTGGAGGTCCTGAAGACAAAGTCATTTGGCAGTTCATTTGATGGAGAG 420
DB	
960	TCCTTCACAGCAATTTGGAGGTCCTGAAGACAAAGTCATTTGGCAGTTCATTTGATGGAGAG 1019
QY	421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 480
DB	
1020	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1079
QY	481 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
DB	
1080	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY	541 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 600
DB	
1140	CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY	601 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAAGATGAAGA 660
DB	
1200	TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAAGATGAAGA 1259
QY	661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 720
DB	
1260	AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 1319
QY	721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
DB	
1320	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
QY	781 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAAAGGAAATGGAGGAAGA 840
DB	
1380	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACAAAGGAAATGGAGGAAGA 1439
QY	841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCA 900
DB	
1440	GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCA 1499
QY	901 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 960

DB	1500 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTAGT 1559
QY	961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 1020
DB	
1560	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
QY	1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTTCTTTTACAAGACAC 1080
DB	
1620	AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTTCTTTTACAAGACAC 1679
QY	1081 TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCCTGCTTACAGA 1140
DB	
1680	TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCCCTGCTTACAGA 1739
QY	1141 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
DB	
1740	AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1799
QY	1201 CTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1260
DB	
1800	CTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1859
QY	1261 TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
DB	
1860	TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1919
QY	1321 AGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAGTG 1380
DB	
1920	AGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAGTG 1979
QY	1381 TGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAGTG 1440
DB	
1980	TGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAGTG 2039
QY	1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
DB	
2040	GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2099
QY	1501 ATTAAGCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGT 1560
DB	
2100	ATTAAGCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGT 2159
QY	1561 ACATAGGCGCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATAGTACTCTTGA 1620
DB	
2160	ACATAGGCGCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATAGTACTCTTGA 2219
QY	1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACTCTACCAGGA 1680
DB	
2220	GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACTCTACCAGGA 2279
QY	1681 GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAAGCA 1740
DB	
2280	GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAAGCA 2339
QY	1741 GGCTGAGGAGGTCAATACCTGAGTGGGAAAAAATTGAACTGCACTCCGCTGACTGGCAGAG 1800
DB	
2340	GGCTGAGGAGGTCAATACCTGAGTGGGAAAAAATTGAACTGCACTCCGCTGACTGGCAGAG 2399
QY	1801 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACCGGATGAGCTGGA 1860
DB	
2400	AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACCGGATGAGCTGGA 2459
QY	1861 CCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCT 1920
DB	
2460	CCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCT 2519
QY	1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCC 1980
DB	
2520	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCC 2579
QY	1981 TCTGAAAGAGAACGTCAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2040
DB	
2580	TCTGAAAGAGAACGTCAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2639

2y 2041 TCAGTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
2b 2640 TCAGTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2699
2y 2101 GCAGGTGGCGCTCGAGGACCGAGTCAGGTCAGCTGCAATGAAGCCACAGGACTTTGGTCC 2160
2b 2700 GCAGGTGGCGCTCGAGGACCGAGTCAGGTCAGCTGCAATGAAGCCACAGGACTTTGGTCC 2759
2y 2161 AGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGTCCTCGGAGAGAGCCATCTCGCC 2220
2b 2760 AGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGTCCTCGGAGAGAGCCATCTCGCC 2819
2y 2221 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAACAACTTGTCTGGACCATCCCAA 2280
2b 2820 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAACAACTTGTCTGGACCATCCCAA 2879
2y 2281 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTGAGATTCTCAGCTTATAG 2340
2b 2880 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTGAGATTCTCAGCTTATAG 2939
2y 2341 GACTGCCATGAACCTCCGAAGACTGCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
2b 2940 GACTGCCATGAACCTCCGAAGACTGCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2999
2y 2401 A 2401
2b 3000 A 3000

RESULT 2
AAD37257
ID AAD37257 standard; DNA; 4825 BP.
AC AAD37257;
XX 21-AUG-2002 (first entry)
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX WO200183695-A2.
XX 08-NOV-2001.
PD 27-APR-2001; 2001WO-US013677.
XX 28-APR-2000; 2000US-020077P.
PR (XIAO/) XIAO X.
XX Xiao X;
PI WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 61-62; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;
Query Match 100.0%; Score 2401; DB 6; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1416
QY 61 TGTGTATACCCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACCTCTTCCA 120
Db 1417 TGTGTATACCCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACCTCTTCCA 1476
QY 121 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGAATGTTGCCAAGGCC 180
Db 1477 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGAATGTTGCCAAGGCC 1536
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAATATGCATATTCTCAACA 240
Db 1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAATATGCATATTCTCAACA 1596
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
Db 1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656
QY 301 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACAGGAGCCCATTT 360
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACAGGAGCCCATTT 1716
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 420
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
QY 421 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
Db 1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
QY 481 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGAAGTGGTGAAGA 540
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGAAGTGGTGAAGA 1896
QY 541 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 600
Db 1897 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1956
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 661 AACTGAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2017 AACTGAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 721 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
Db 2077 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAGGAAATGGAGGAAGA 840
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAGGAAATGGAGGAAGA 2196
QY 841 GCCTCTTGAGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAGGTCTTCA 900

Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256
QY 901 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTCTCTCACTCAATGCTGGTGGTGTAGT 960
Db 2257 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTCTCTCACTCAATGCTGGTGGTGTAGT 2316
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 1020
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 2376
QY 1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAC 1080
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAC 2436
QY 1081 TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCTGCTGGCTTACAGA 1140
Db 2437 TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCTGCTGGCTTACAGA 2496
QY 1141 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 1200
Db 2497 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 2556
QY 1201 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
Db 2557 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616
QY 1261 TCACACAGATGTTTATCAACAACCTGGATGAAACAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
Db 2617 TCACACAGATGTTTATCAACAACCTGGATGAAACAACAGCCAAAAAATCCTGAGATCCCTGGA 2676
QY 1321 AGGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGATAACATCAAACTTCAAGTGGAG 1380
Db 2677 AGGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGATAACATCAAACTTCAAGTGGAG 2736
QY 1381 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
Db 2737 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2796
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 2797 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2856
QY 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACITTTCCAGCAGTTTCAGAGCAGAACGATGT 1560
Db 2857 ATTAAGCCGGCAGGCACCTATTGGAGGCGACITTTCCAGCAGTTTCAGAGCAGAACGATGT 2916
QY 1561 ACATAGGGCCTTCAAGAGGGAATTGAAACCTAAGAACCTGTAATCATGAGTACTCTTGA 1620
Db 2917 ACATAGGGCCTTCAAGAGGGAATTGAAACCTAAGAACCTGTAATCATGAGTACTCTTGA 2976
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGGA 1680
Db 2977 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGGA 3036
QY 1681 GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
Db 3037 GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 3096
QY 1741 GGTGTAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
Db 3097 GGTGTAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 3156
QY 1801 AAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1860
Db 3157 AAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 3216
QY 1861 CCTCAAGCTGCGCCAAAGCTCAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCT 1920
Db 3217 CCTCAAGCTGCGCCAAAGCTCAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCT 3276
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTGCGCC 1980
Db 3277 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTGCGCC 3336

RESULT 3

AAD37263

ID AAD37263 standard; DNA; 4848 BP.

XX

AC AAD37263;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.

OS Cytomegalovirus.

OS Unidentified.

OS Chimeric.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

PT gene.

XX

PS Example 1; Page 68-70; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX

SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 2401; DB 6; Length 4848;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	60
Db	1380	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	1439
QY	61	TGTTGATACCACTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA	120
Db	1440	TGTTGATACCACTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA	1499
QY	121	AGTTTGGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC	180
Db	1500	AGTTTGGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC	1559
QY	181	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA	240
Db	1560	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA	1619
QY	241	GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA	300
Db	1620	GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA	1679
QY	301	GAGCTATGCTTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT	360
Db	1680	GAGCTATGCTTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT	1739
QY	361	TCCTTCACAGCATTTGGAAGCTCCTGAAACAAAGTCAATTTGGCAGTTTCAATGATGGAGAG	420
Db	1740	TCCTTCACAGCATTTGGAAGCTCCTGAAACAAAGTCAATTTGGCAGTTTCAATGATGGAGAG	1799
QY	421	TGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGATTTATCGTGGCTTCTTTTC	480
Db	1800	TGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGATTTATCGTGGCTTCTTTTC	1859
QY	481	TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	540
Db	1860	TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1919
QY	541	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG	600
Db	1920	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG	1979
QY	601	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA	660
Db	1980	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA	2039
QY	661	AAGTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC	720
Db	2040	AAGTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC	2099
QY	721	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT	780
Db	2100	TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT	2159
QY	781	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA	840

Db	2160	GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA	2219
QY	841	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAACAACATAAGGTGCTTCA	900
Db	2220	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAACAACATAAGGTGCTTCA	2279
QY	901	AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT	960
Db	2280	AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT	2339
QY	961	TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG	1020
Db	2340	TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG	2399
QY	1021	AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAC	1080
Db	2400	AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAC	2459
QY	1081	TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA	1140
Db	2460	TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA	2519
QY	1141	AGCTGAAACAACTGCCAATGTCTTACAGGATGTACCCGTAAAGAAAGCTCCTAGAAGA	1200
Db	2520	AGCTGAAACAACTGCCAATGTCTTACAGGATGTACCCGTAAAGAAAGCTCCTAGAAGA	2579
QY	1201	CTCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAGGTGAAATGAAGC	1260
Db	2580	CTCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAGGTGAAATGAAGC	2639
QY	1261	TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGA	1320
Db	2640	TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGA	2699
QY	1321	AGGTTCCGATGATGACGCTCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG	1380
Db	2700	AGGTTCCGATGATGACGCTCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG	2759
QY	1381	TGAACTTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAAGCCAGTTCTGACCAGTG	1440
Db	2760	TGAACTTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAAGCCAGTTCTGACCAGTG	2819
QY	1441	GAAAGCTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGA	1500
Db	2820	GAAAGCTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGA	2879
QY	1501	ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAGAACGAGAACGATGT	1560
Db	2880	ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAGAACGAGAACGATGT	2939
QY	1561	ACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTATATCATGATGACTCTTGA	1620
Db	2940	ACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTATATCATGATGACTCTTGA	2999
QY	1621	GACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACTCTACCAGGA	1680
Db	3000	GACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACTCTACCAGGA	3059
QY	1681	GCCCAGAGAGCTGCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAGCA	1740
Db	3060	GCCCAGAGAGCTGCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAGCA	3119
QY	1741	GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	1800
Db	3120	GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	3179
QY	1801	AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA	1860
Db	3180	AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA	3239
QY	1861	CCTCAAGCTGGCCAGCTGAGGTGATCAAGGGATCTTGGAGCCCGTGGCGGATCTCCT	1920

Db 3240 CCTAAGCTGGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT 3299
Qy 1921 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTGCGCC 1980
Db 3300 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTGCGCC 3359
Qy 1981 TCTGAAAGAGAAACGTGAGCCACGTCAATGACTTGTCTGCCAGCTTACCACCTTTGGGCAT 2040
Db 3360 TCTGAAAGAGAAACGTGAGCCACGTCAATGACTTGTCTGCCAGCTTACCACCTTTGGGCAT 3419
Qy 2041 TCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGGAACACAGATGGAAGCTTCT 2100
Db 3420 TCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGGAACACAGATGGAAGCTTCT 3479
Qy 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTTGGTCC 2160
Db 3480 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTTGGTCC 3539
Qy 2161 AGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
Db 3540 AGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3599
Qy 2221 AAACAAAGTGCCTACTATATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAA 2280
Db 3600 AAACAAAGTGCCTACTATATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAA 3659
Qy 2281 AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTAGATCTCTCAGCTTATAG 2340
Db 3660 AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTAGATCTCTCAGCTTATAG 3719
Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 2400
Db 3720 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 3779
Qy 2401 A 2401
Db 3780 A 3780

RESULT 4
AAD37264

ID AAD37264 standard; DNA; 5060 BP.

AC AAD37264;

DT 21-AUG-2002 (first entry)

DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.

XX WO200183695-A2.

PN 08-NOV-2001.

PF 27-APR-2001; 2001WO-US013677.

PR 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

PA Xiao X;

DR WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.

XX Example 1; Page 70-71; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
CC small polyA signal sequence

XX Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Qy Query Match 100.0%; Score 2401; DB 6; Length 5060;
Db Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 1592 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1651

Qy 61 TGTGATACACCTATCCAGATAAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
Db 1652 TGTGATACACCTATCCAGATAAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 1711

Qy 121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 1712 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1771

Qy 181 ACCTAAAGTGACTAAAGAGAACAATTTTCAGTTACATCATCAAAATGCATCTTCAACA 240
Db 1772 ACCTAAAGTGACTAAAGAGAACAATTTTCAGTTACATCATCAAAATGCATCTTCAACA 1831

Qy 241 GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
Db 1832 GATCAGCGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1891

Qy 301 GAGCTATGCCTACACACAGGCTGCTTATGTACACCTCTGACCCCTACACGGAGCCCAT 360
Db 1892 GAGCTATGCCTACACACAGGCTGCTTATGTACACCTCTGACCCCTACACGGAGCCCAT 1951

Qy 361 TCCITCACAGCATTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTTATGATGGAGAG 420
Db 1952 TCCITCACAGCATTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTTATGATGGAGAG 2011

Qy 421 TGAAGTAAACCTGGACCGTTATCAACAGCCTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
Db 2012 TGAAGTAAACCTGGACCGTTATCAACAGCCTTAGAAGAAGTATTATCGTGGCTTCTTTC 2071

Qy 481 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 540
Db 2072 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 2131

Qy 541 CCAGTTTCATCTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGTTGG 600
Db 2132 CCAGTTTCATCTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGTTGG 2191

Qy 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 660
Db 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 2251

Qy 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2311

DR WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;
Query Match 93.7%; Score 2250; DB 6; Length 3999;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 659
61 TGTGTATACCACTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
660 TGTGTATACCACTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 719
121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
720 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATCTCAACA 240
780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATCTCAACA 839
241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 300
840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 899
301 GAGCTATGCCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 360
900 GAGCTATGCCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 959
361 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAAGTCATTTGGCAGTTTCATTGTATGGAGAG 420
960 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAAGTCATTTGGCAGTTTCATTGTATGGAGAG 1019
421 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
1020 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
481 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
1080 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
541 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 600
1140 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
601 TAATATTTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 660
1200 TAATATTTCTACAAATGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1259

QY 661 AACTGAAGTACAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGTTAGC 720
Db 1260 AACTGAAGTACAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGTTAGC 1319
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 780
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAATCAGAAACT 1379
QY 781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 840
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1439
QY 841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAACGCCAAGTACAAACACATAAGGTGCTTCA 900
Db 1440 GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAACGCCAAGTACAAACACATAAGGTGCTTCA 1499
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 960
Db 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1559
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1020
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1619
QY 1021 AGATCGATGGCAAAACATCTGTAGATGGACAGAACGCCGCTGGGTTCTTTTACAAGAC-- 1078
Db 1620 AGATCGATGGCAAAACATCTGTAGATGGACAGAACGCCGCTGGGTTCTTTTACAAGACCA 1679
QY 1079 ----- 1078
Db 1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 1739
QY 1079 ----- 1078
Db 1740 GGTGACACAACTGTGGTTACTAAGGAACTGCCATCTCCAAACTAGAAATGCCATCTTC 1799
QY 1079 -----ACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAA 1119
Db 1800 CTTGATGTTGGAGTACCTACTCTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAA 1859
QY 1120 GTTCTTTCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCG 1179
Db 1860 GTTCTTTCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCG 1919
QY 1180 TAAGGAAAGGCTCCTAGAAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAGCCA 1239
Db 1920 TAAGGAAAGGCTCCTAGAAAGCTCCAAAGGAGTAAAGAGCTGATGAAACAACTGGCAGA 1979
QY 1240 CCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAGCCA 1299
Db 1980 CCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACAGCCA 2039
QY 1300 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGA 1359
Db 2040 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGA 2099
QY 1360 TAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 1419
Db 2100 TAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 2159
QY 1420 GGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGTGTG 1479
Db 2160 GGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGTGTG 2219
QY 1480 GGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 1539
Db 2220 GGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 2279
QY 1540 AGTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGGGGAATGAAAACTAAAGAAC 1599
Db 2280 AGTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGGGGAATGAAAACTAAAGAAC 2339
QY 1600 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGG 1659

Db 2340 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGG 2399
QY 1660 ACTAGAGAAACTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGT 1719
Db 2400 ACTAGAGAAACTCTACCAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGT 2459
QY 1720 CACTCGGCTTCTACGAAAGCAGGCTGAGAGAGTCAATACTAGTGGGAAAAATTTGAACCT 1779
Db 2460 CACTCGGCTTCTACGAAAGCAGGCTGAGAGAGTCAATACTAGTGGGAAAAATTTGAACCT 2519
QY 1780 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCA 1839
Db 2520 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCA 2579
QY 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGCTGATCAAGGGATCCTG 1899
Db 2580 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGCTGATCAAGGGATCCTG 2639
QY 1900 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATACCTCGAGAAAGTCAAGGC 1959
Db 2640 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATACCTCGAGAAAGTCAAGGC 2699
QY 1960 ACTTCGAGGAGAAATTCGCCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGTCTG 2019
Db 2700 ACTTCGAGGAGAAATTCGCCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGTCTG 2759
QY 2020 CCAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 2079
Db 2760 CCAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 2819
QY 2080 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2139
Db 2820 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2879
QY 2140 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTGTCAGGGTCC 2199
Db 2880 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTGTCAGGGTCC 2939
QY 2200 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGACTCAAAAC 2259
Db 2940 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGACTCAAAAC 2999
QY 2260 AACTTGTGGGACCATCCCAAATGACAGAGCTCTACCAGCTTTTAGCTGACCTGAATAA 2319
Db 3000 AACTTGTGGGACCATCCCAAATGACAGAGCTCTACCAGCTTTTAGCTGACCTGAATAA 3059
QY 2320 TGTGAGATTCTCAGCTTATAGGACTGGCATGAAACTCCGAGAGACTGCAGAAGGCCCTTTG 2379
Db 3060 TGTGAGATTCTCAGCTTATAGGACTGGCATGAAACTCCGAGAGACTGCAGAAGGCCCTTTG 3119
QY 2380 CTTGGATCTCTTGAGCCTGTCA 2401
Db 3120 CTTGGATCTCTTGAGCCTGTCA 3141

RESULT 6

AAD37256
ID AAD37256 standard; DNA; 4966 BP.

AC AAD37256;

KX 21-AUG-2002 (first entry)

QY Adeno-associated virus vector plasmid, AAV-MCK-delta3990.

QW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
QW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
QW Becker muscular dystrophy; ds.

QX Homo sapiens.

QX Unidentified.

QX Chimeric.

XX WO200183695-A2.
FN
XX 08-NOV-2001.
PD
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
FI Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 59-60; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match 93.7%; Score 2250; DB 6; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1416
QY 61 TGTGTATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
Db 1417 TGTGTATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476
QY 121 AGTTTTCCTCAACAAGTGAAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 1477 AGTTTTCCTCAACAAGTGAAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTTCAGTTACATCATCAATGCACTATTCTCAACA 240
Db 1537 ACCTAAAGTGACTAAAGAAGAACATTTTTCAGTTACATCATCAATGCACTATTCTCAACA 1596
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
Db 1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656
QY 301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 360
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 1716
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGGAACAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 420
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGGAACAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 1776
QY 421 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
Db 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

QY 481 TGCTCAGGACACATTGCAAGCACAAAGCAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGCAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 600
Db 1897 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 660
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 661 AACTGAAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2017 AACTGAAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGATGTTTAATGGATCTCCAGAATCAGAAACT 780
Db 2077 TAGCATGGAAAAACAAAGCAATTTACATAGATGTTTAATGGATCTCCAGAATCAGAAACT 2136
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 840
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 2196
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAGGTGCTTCA 900
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 960
Db 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 2316
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAACCGTGGGTTCTTTTACAAGAC-- 1078
Db 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAAACCGTGGGTTCTTTTACAAGACCA 2436
QY 1079 ----- 1078
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCCTCTCCTACTCAGACTGTTACTCT 2496
QY 1079 ----- 1078
Db 2497 GGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 2556
QY 1079 -----ACTCATAGATTACTGCAACAGTTCCCTCGTGGACCTGGAAAA 1119
Db 2557 CTTGATGTTGAGGTACCTTACTCATAGATTACTGCAACAGTTCCCTCGTGGACCTGGAAAA 2616
QY 1120 GTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 1179
Db 2617 GTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 2676
QY 1180 TAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 1239
Db 2677 TAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 2736
QY 1240 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCA 1299
Db 2737 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCA 2796
QY 1300 AAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGGA 1359
Db 2797 AAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGGA 2856
QY 1360 TAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 1419
Db 2857 TAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 2916
QY 1420 GGAAGCCAGTTCTGACCCAGTGGAAAGCGTCTGCACTTTCTCTGCAAGAACTTCTTGGTGTG 1479

RESULT 7
AAD37262
ID AAD37262 standard; DNA; 4990 BP.
XX
AC AAD37262;

Db 2917 GGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 2976
QY 1480 GCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCACTTTCCAGC 1539
Db 2977 GCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCACTTTCCAGC 3036
QY 1540 AGTTCAGAAGCAGAACGATGATGATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACC 1599
Db 3037 AGTTCAGAAGCAGAACGATGATGATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACC 3096
QY 1600 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGG 1659
Db 3097 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGG 3156
QY 1660 ACTAGAGAAAACTCTACAGGAGCCCCAGAGAGCTGCTCCTGAGGAGAGAGGCCAGAAATGT 1719
Db 3157 ACTAGAGAAAACTCTACAGGAGCCCCAGAGAGCTGCTCCTGAGGAGAGAGGCCAGAAATGT 3216
QY 1720 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 1779
Db 3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 3276
QY 1780 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCA 1839
Db 3277 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCA 3336
QY 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTG 1899
Db 3337 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTG 3396
QY 1900 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1959
Db 3397 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 3456
QY 1960 ACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGTCTCG 2019
Db 3457 ACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGTCTCG 3516
QY 2020 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACCTTGGAAAGACCT 2079
Db 3517 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACCTTGGAAAGACCT 3576
QY 2080 GAACACCAGATGGAAGCTTTGTCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2139
Db 3577 GAACACCAGATGGAAGCTTTGTCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 3636
QY 2140 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGGTCC 2199
Db 3637 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGGTCC 3696
QY 2200 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAC 2259
Db 3697 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAC 3756
QY 2260 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 2319
Db 3757 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3816
QY 2320 TGTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCCCTTTG 2379
Db 3817 TGTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCCCTTTG 3876
QY 2380 CTTGGATCTCTTGAGCCTGTCA 2401
Db 3877 CTTGGATCTCTTGAGCCTGTCA 3898

XX 21-AUG-2002 (first entry)
DT Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
KW
XX
OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX
XX WO200183695-A2.
PN 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US013677.
PF 28-APR-2000; 2000US-0200777P.
XX
XX (XIAO/) XIAO X.
PA Xiao X;
XX
PI WPI; 2002-049342/06.
XX
DR New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
XX
XX Example 1; Page 67-68; 71pp; English.
PS
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a
XX cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;
Query Match 93.7%; Score 2250; DB 6; Length 4990;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
2Y 1 ACATGCATTCAACATCGCCAGATATCAATAGGCATAGAGAACTACTCGATCCTGAAGA 60
2b 1381 ACATGCATTCAACATCGCCAGATATCAATAGGCATAGAGAACTACTCGATCCTGAAGA 1440
2Y 61 TGTGTATACACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
2b 1441 TGTGTATACACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1500
2Y 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
2b 1501 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1560
2Y 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 240
2b 1561 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 1620
2Y 241 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 300
2b 1621 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 1680

QY 301 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGTACCCCTACACGAGCCCCATT 360
Db 1681 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGTACCCCTACACGAGCCCCATT 1740
QY 361 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 420
Db 1741 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1800
QY 421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 480
Db 1801 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1860
QY 481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1861 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1920
QY 541 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCCGGTTGG 600
Db 1921 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCCGGTTGG 1980
QY 601 TAATATTCTAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 660
Db 1981 TAATATTCTAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2040
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2041 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2100
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
Db 2101 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2160
QY 781 GAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGAAAAATGGAGGAAGA 840
Db 2161 GAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGAAAAATGGAGGAAGA 2220
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAGGTGCTTCA 900
Db 2221 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAGGTGCTTCA 2280
QY 901 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 960
Db 2281 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2340
QY 961 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG 1020
Db 2341 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG 2400
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGTCTTTTACAAAGAC-- 1078
Db 2401 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGTCTTTTACAAAGACCA 2460
QY 1079 ----- 1078
Db 2461 GCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCTCTACTCAGACTGTTACTCT 2520
QY 1079 ----- 1078
Db 2521 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTTC 2580
QY 1079 -----ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAA 1119
Db 2581 CTTGATGTTGGAGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAA 2640
QY 1120 GTTTCCTTGCCTGGTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 1179
Db 2641 GTTTCCTTGCCTGGTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 2700
QY 1180 TAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 1239
Db 2701 TAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 2760

QY 1240 CCTCAAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCA 1299
Db 2761 CCTCAAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCA 2820
QY 1300 AAAAATCCTGAGATCCCTGGAAGTTTCCGATGATGCAGTCCCTGTACAAAAGACGTTTGA 1359
Db 2821 AAAAATCCTGAGATCCCTGGAAGTTTCCGATGATGCAGTCCCTGTACAAAAGACGTTTGA 2880
QY 1360 TAACATGAACCTTCAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCCCATTT 1419
Db 2881 TAACATGAACCTTCAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCCCATTT 2940
QY 1420 GGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTG 1479
Db 2941 GGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTG 3000
QY 1480 GCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 1539
Db 3001 GCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 3060
QY 1540 AGTTTCAGAACGACGATGTACATAGGGCCCTCAAGAGGGGAATTGAAAACCTAAAGAAC 1599
Db 3061 AGTTTCAGAACGACGATGTACATAGGGCCCTCAAGAGGGGAATTGAAAACCTAAAGAAC 3120
QY 1600 TGTAAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGG 1659
Db 3121 TGTAAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGG 3180
QY 1660 ACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGATGT 1719
Db 3181 ACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCTCCTGAGGAGAGAGCCCGAGATGT 3240
QY 1720 CACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCT 1779
Db 3241 CACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCT 3300
QY 1780 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCA 1839
Db 3301 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCA 3360
QY 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGCATCAAGGGATCCTG 1899
Db 3361 AGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGCATCAAGGGATCCTG 3420
QY 1900 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGC 1959
Db 3421 GCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGC 3480
QY 1960 ACTTCGAGGAGAAATTGGGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 2019
Db 3481 ACTTCGAGGAGAAATTGGGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 3540
QY 2020 CCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 2079
Db 3541 CCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 3600
QY 2080 GAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2139
Db 3601 GAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 3660
QY 2140 AGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 2199
Db 3661 AGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 3720
QY 2200 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGACTCAAAC 2259
Db 3721 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGACTCAAAC 3780
QY 2260 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAA 2319
Db 3781 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAA 3840
QY 2320 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGAGGCCCTTTG 2379

Db 3841 TGTGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAGAAGCCCTTTG 3900
QY 2380 CTTGGATCTCTTGAGCCTGTCA 2401
Db 3901 CTTGGATCTCTTGAGCCTGTCA 3922
RESULT 8
AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX AAD37230;
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 43-44; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
Query Match 86.1%; Score 2067; DB 6; Length 4182;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
QY 1 ACATGCATTCAACATCGGCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
Db 600 ACATGCATTCAACATCGGCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659
QY 61 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCCTCTTCCA 120
Db 660 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCCTCTTCCA 719

121 AGTTTGGCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
122 |||||
720 AGTTTGGCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
721 |||||
181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA 240
182 |||||
780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA 839
781 |||||
241 GATCAGGTGAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 300
242 |||||
840 GATCAGGTGAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 899
841 |||||
301 GAGCTATGCTTACACACAGGCTGTTATGTCAACCCTCTGACCCCTACACGAGCCCCATT 360
302 |||||
900 GAGCTATGCTTACACACAGGCTGTTATGTCAACCCTCTGACCCCTACACGAGCCCCATT 959
901 |||||
361 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 420
362 |||||
960 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1019
961 |||||
421 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 480
422 |||||
1020 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 1079
1021 |||||
481 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
482 |||||
1080 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
1081 |||||
541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGTTGG 600
542 |||||
1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGTTGG 1199
1141 |||||
601 TAATATTCTAACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 660
602 |||||
1200 TAATATTCTAACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1259
1201 |||||
661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 720
662 |||||
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 1319
1261 |||||
721 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 780
722 |||||
1320 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 1379
1321 |||||
781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 840
782 |||||
1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1439
1381 |||||
841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAAATAGGTGCTTCA 900
842 |||||
1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAAATAGGTGCTTCA 1499
1441 |||||
901 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 960
902 |||||
1500 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 1559
1501 |||||
961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 1020
962 |||||
1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 1619
1561 |||||
1021 AGATCGATGGGCACAAATCTGTAGATGGACAGAAGACCGTGGTTCTTTTACAAGAC-- 1078
1022 |||||
1620 AGATCGATGGGCACAAATCTGTAGATGGACAGAAGACCGTGGTTCTTTTACAAGACAT 1679
1621 |||||
1079 ----- 1078
1680 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1739
1681 |||||
1079 ----- 1078
1740 AAAAGAAAGATGCAGTGAACAAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1799
1741 |||||
1079 ----- 1078

Db 1800 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 1859
Qy 1079 ----- 1078
Db 1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1919
Qy 1079 ----- 1078
Db 1920 CCAGAAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1979
Qy 1079 ----- 1078
Db 1980 ACTTGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGATTTCCCTCTGGACCTGGA 2039
Qy 1117 AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1176
Db 2040 AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 2099
Qy 1177 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1236
Db 2100 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2159
Qy 1237 AGACCTCCAAGTGAAAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAAG 1296
Db 2160 AGACCTCCAAGTGAAAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAAG 2219
Qy 1297 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCA 1356
Db 2220 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCA 2279
Qy 1357 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA 1416
Db 2280 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA 2339
Qy 1417 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTTCTGTT 1476
Db 2340 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTTCTGTT 2399
Qy 1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 1536
Db 2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 2459
Qy 1537 AGCAGTTCAGAAAGCAGAACGATGTATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1596
Db 2460 AGCAGTTCAGAAAGCAGAACGATGTATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2519
Qy 1597 ACCTGTAAATCATGAGTACTCTTGGAGTGTACGAATATTTCTGACAGAGCAGCCCTTGG 1656
Db 2520 ACCTGTAAATCATGAGTACTCTTGGAGTGTACGAATATTTCTGACAGAGCAGCCCTTGG 2579
Qy 1657 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCCGAGAA 1716
Db 2580 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCCGAGAA 2639
Qy 1717 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1776
Db 2640 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 2699
Qy 1777 CTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 1836
Db 2700 CTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 2759
Qy 1837 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGTGAAGTGAATCAAGGGATC 1896
Db 2760 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGTGAAGTGAATCAAGGGATC 2819
Qy 1897 CTGGCAGCCCGTGGCGGATCTCCTCAATGACTCTCTCCAAGATCACCTCGAGAAAAAGTCAA 1956
Db 2820 CTGGCAGCCCGTGGCGGATCTCCTCAATGACTCTCTCCAAGATCACCTCGAGAAAAAGTCAA 2879
Qy 1957 GGCACCTCGAGAGAAAAATTGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGC 2016
Db 1957 GGCACCTCGAGAGAAAAATTGCGCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGC 2016

Db 2880 GGCACCTCGAGGAGAAATTGGCCTCTGAAGAGAACGTTGAGCCACGTCATGACCTTGC 2939
Qy 2017 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 2076
Db 2940 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 2999
Qy 2077 CCTGAACACCAGATGGAAGCTTCTGACGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCA 2136
Db 3000 CCTGAACACCAGATGGAAGCTTCTGACGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCA 3059
Qy 2137 TGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTGTCAGGG 2196
Db 3060 TGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTGTCAGGG 3119
Qy 2197 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCA 2256
Db 3120 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCA 3179
Qy 2257 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 2316
Db 3180 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 3239
Qy 2317 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAGAAAGGCCCT 2376
Db 3240 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAGAAAGGCCCT 3299
Qy 2377 TTGCTTGGATCTCTTGAGCCTGTCA 2401
Db 3300 TTGCTTGGATCTCTTGAGCCTGTCA 3324

RESULT 9

AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-020077P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX

SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match 86.1%; Score 2067; DB 6; Length 5149;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

Qy 1 ACATGCAATTCACATCGCCAGATATCAATTAGSCATAGAGAACTACTCGATCCTGAAGA 60
Db 1357 ACATGCAATTCACATCGCCAGATATCAATTAGSCATAGAGAACTACTCGATCCTGAAGA 1416
Qy 61 TGTGATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCATCACTCTTCCA 120
Db 1417 TGTGATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCATCACTCTTCCA 1476
Qy 121 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 1477 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536
Qy 181 ACCTAAAGTGACTAAAGAAACAATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 240
Db 1537 ACCTAAAGTGACTAAAGAAACAATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 1596
Qy 241 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTTCCCTTAAGCTTCGATTCAA 300
Db 1597 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTTCCCTTAAGCTTCGATTCAA 1656
Qy 301 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCTCTGACCCCTACACGGAGCCCAT 360
Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCTCTGACCCCTACACGGAGCCCAT 1716
Qy 361 TCCTTCACAGCATTTGGAAGTCTCTGAAGACAAAGTCAATTTGGCAGTTTATGATGGAGAG 420
Db 1717 TCCTTCACAGCATTTGGAAGTCTCTGAAGACAAAGTCAATTTGGCAGTTTATGATGGAGAG 1776
Qy 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 480
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1836
Qy 481 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
Qy 541 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 600
Db 1897 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1956
Qy 601 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 660
Db 1957 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 2016
Qy 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 720
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076
Qy 721 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
Db 2077 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
Qy 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAAGAAATGAGGAAGA 840
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAAGAAATGAGGAAGA 2196
Qy 841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAGGTGCTTCA 900

Db 2197 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTACAAACAATAAGGTGCTTCA 2256
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGCTCAATTCTCTCACTCACATGGTGGTGGTAGT 960
Db 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGCTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
QY 1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGAC-- 1078
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 2436
QY 1079 ----- 1078
Db 2437 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTTCAGA 2496
QY 1079 ----- 1078
Db 2497 AAAAGAAGATGCAGTGAACAAGATTACACAACACTGGCTTTTAAGATCAAAATGAAATGTT 2556
QY 1079 ----- 1078
Db 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2616
QY 1079 ----- 1078
Db 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2676
QY 1079 ----- 1078
Db 2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAA 2736
QY 1079 -----ACTCATAGATTACTGCAACAGATTCCCCCTGGACCTGGA 1116
Db 2737 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2796
QY 1117 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTAC 1176
Db 2797 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTAC 2856
QY 1177 CCGTAAGGAAAGGTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 1236
Db 2857 CCGTAAGGAAAGGTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCA 2916
QY 1237 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAACAACTGGATGAACAACAG 1296
Db 2917 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAACAACTGGATGAACAACAG 2976
QY 1297 CCAAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTCTGTTACAAAGACGTTT 1356
Db 2977 CCAAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTCTGTTACAAAGACGTTT 3036
QY 1357 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 1416
Db 3037 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 3096
QY 1417 TTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1476
Db 3097 TTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 3156
QY 1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTTATGGAGGCGACTTCC 1536
Db 3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTTATGGAGGCGACTTCC 3216
QY 1537 AGCAGTTCAGAACGACGATGTACATAGGCTTCAAGAGGGAATTGAAACTAAAGA 1596
Db 3217 AGCAGTTCAGAACGACGATGTACATAGGCTTCAAGAGGGAATTGAAACTAAAGA 3276
QY 1597 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCCTTTGGA 1656

Db 3277 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGCACAGAGCAGCCTTTTGA 3336
QY 1657 AGGACTAGAGAAACTCTACACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1716
Db 3337 AGGACTAGAGAAACTCTACACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3396
QY 1717 TGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1776
Db 3397 TGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3456
QY 1777 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 1836
Db 3457 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAAC 3516
QY 1837 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATC 1896
Db 3517 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATC 3576
QY 1897 CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 1956
Db 3577 CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 3636
QY 1957 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCTTGC 2016
Db 3637 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAACGTCGAGCCACGTCATGACCTTGC 3696
QY 2017 TCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2076
Db 3697 TCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3756
QY 2077 CCGTAAACACAGATGGAAGCTTCTGAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCA 2136
Db 3757 CCGTAAACACAGATGGAAGCTTCTGAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCA 3816
QY 2137 TGAAGCCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGG 2196
Db 3817 TGAAGCCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGG 3876
QY 2197 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCCTACTATATCAACCCAGAGACTCA 2256
Db 3877 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCCTACTATATCAACCCAGAGACTCA 3936
QY 2257 AACAACCTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 2316
Db 3937 AACAACCTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 3996
QY 2317 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAGGCCCT 2376
Db 3997 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAGGCCCT 4056
QY 2377 TTGCTTGGATCTCTTGAGCCTGTCA 2401
Db 4057 TTGCTTGGATCTCTTGAGCCTGTCA 4081

RESULT 10

ABK81999

ID ABK81999 standard; DNA; 5462 BP.

XX

AC ABK81999;

XX

DT 13-AUG-2002 (first entry)

XX

DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX

KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200229056-A2.

XX

PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
PS WPI; 2002-435334/46.
XX
DR
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Disclosure; Fig 14; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
CC invention
XX
SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;
Query Match 76.6%; Score 1839.4; DB 6; Length 5462;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 76; Indels 210; Gaps 2;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
DB 799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 858
QY 61 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTTCCA 120
DB 859 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTTCCA 918
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
DB 919 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 978
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACATTTCTCAACA 240
DB 979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACATTTCTCAACA 1038
QY 241 GATCAGGTGCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 300
DB 1039 GATCAGGTGCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 1098
QY 301 GAGCTATGCCTACACAGGGCTGCTTATGTACCCACCTCTGACCCCTACAGGAGCCCAT 360
DB 1099 GAGCTATGCCTACACAGGGCTGCTTATGTACCCACCTCTGACCCCTACAGGAGCCCAT 1158
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGTCAATTTGGCAGTTTCATTGTGAGAG 420
DB 1159 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGTCAATTTGGCAGTTTCATTGTGAGAG 1218
QY 421 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGATTATTCGTGGCTTCTTC 480
DB 1219 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAGATTATTCGTGGCTTCTTC 1278
QY 481 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAAGTGGTGAAGA 540
DB 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAAGTGGTGAAGA 1338
QY 541 CCAGTTTCATCTAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG 600
DB

Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGCGCGGTGG 1398
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 660
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 661 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTTAGC 720
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTTAGC 1518
QY 721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCTGGACTG 1558
QY 781 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAAAGAAATGGAGGAAGA 840
Db 1559 -----AC 1560
QY 841 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAAAAGCCCAAGTACAAACAATAAGGTGCTTCA 900
Db 1561 CACTATTGGAGCCTCTCTCTACTCAGACTGTTACTCTGGTGACACAACCTGTGTTACTAA 1620
QY 901 AGAAGATCTAGAAACAAGAACAAAGTCAGGTCATTAATCTCTCACTCACATGGTGGTGTAGT 960
Db 1621 GGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGTGTTGGAG----- 1669
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 1020
Db 1670 ----- 1669
QY 1021 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
Db 1670 ----- 1669
QY 1081 TCATAGATTACTGCAACAGTTTCCCTCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1140
Db 1670 -CATAGATTACTGCAACAGTTTCCCTCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1728
QY 1141 AGCTGAAACAACCTGCCAATGTCTCTACAGATGCTACCCGTAAAGAAAGGCTCCTTAGAAGA 1200
Db 1729 AGCTGAAACAACCTGCCAATGTCTCTACAGATGCTACCCGTAAAGAAAGGCTCCTTAGAAGA 1788
QY 1201 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
Db 1789 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1848
QY 1261 TCACACAGATGTTTATCAACACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
Db 1849 TCACACAGATGTTTATCAACACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1908
QY 1321 AGTTCCGATGATGAGTCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
Db 1909 AGTTCCGATGATGAGTCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1968
QY 1381 TGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCCAGTTTCTGACCAGTG 1440
Db 1969 TGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCCAGTTTCTGACCAGTG 2028
QY 1441 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 2029 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2088
QY 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1560
Db 2089 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCAGAACGATGT 2148
QY 1561 ACATAGGGCCTTCAAGAGGAATTTGAAACCTTAAAGAAACCTGTAATCATGAGTACTCTTGA 1620
Db 2149 ACATAGGGCCTTCAAGAGGAATTTGAAACCTTAAAGAAACCTGTAATCATGAGTACTCTTGA 2208
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 1680
Db 2209 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 2268

2Y 1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
|||
2b 2269 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2328
|||
2Y 1741 GGCTGAGAGAGTCAATACTAGTGGGAAAAATTGAACTGCACTCCGCTGACTGGCAGAG 1800
|||
2b 2329 GGCTGAGAGAGTCAATACTAGTGGGAAAAATTGAACTGCACTCCGCTGACTGGCAGAG 2388
|||
2Y 1801 AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
|||
2b 2389 AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2448
|||
2Y 1861 CCTCAAGCTCGGCCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCCGTGGCGATCTCCT 1920
|||
2b 2449 CCTCAAGCTCGGCCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCCGTGGCGATCTCCT 2508
|||
2Y 1921 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
|||
2b 2509 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 2568
|||
2Y 1981 TCTGAAAGAGAAAGTGAAGCCAGTCAATGACCTTGTCCGACGTTTACCCTTTGGGCAT 2040
|||
2b 2569 TCTGAAAGAGAAAGTGAAGCCAGTCAATGACCTTGTCCGACGTTTACCCTTTGGGCAT 2628
|||
2Y 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 2100
|||
2b 2629 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 2688
|||
2Y 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGCTGATGAAGCCACAGGGACTTTGGTCC 2160
|||
2b 2689 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGCTGATGAAGCCACAGGGACTTTGGTCC 2748
|||
2Y 2161 AGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCC 2220
|||
2b 2749 AGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCC 2808
|||
2Y 2221 AAACAAAGTGCCTACTATATCAACACGAGACTCAAAACAATTTGCTGGGACCATCCCA 2280
|||
2b 2809 AAACAAAGTGCCTACTATATCAACACGAGACTCAAAACAATTTGCTGGGACCATCCCA 2868
|||
2Y 2281 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAG 2340
|||
2b 2869 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAG 2928
|||
2Y 2341 GACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGTTGGATCTCTTGAGCCTGTC 2400
|||
2b 2929 GACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGTTGGATCTCTTGAGCCTGTC 2988
|||
2Y 2401 A 2401
|
2b 2989 A 2989

RESULT 11
AAD37238
ID AAD37238 standard; DNA; 3531 BP.
XX
AC AAD37238;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
DS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.
PF
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.
XX
PS Example 1; Page 50-51; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match 72.3%; Score 1737; DB 6; Length 3531;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2074; Conservative 0; Mismatches 0; Indels 327; Gaps 1;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
|||
Db 600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 659
|||
QY 61 TGTGTATACCACTATCCAGATAAGAAAGTCCATCTTAATGTATCATCATCATCTCTTCCA 120
|||
Db 660 TGTGTATACCACTATCCAGATAAGAAAGTCCATCTTAATGTATCATCATCATCTCTTCCA 719
|||
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAGGCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
|||
Db 720 AGTTTTCCTCAACAAGTGAGCATTTGAGGCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
|||
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTTACATCATCAATGCACATATTCTCAACA 240
|||
Db 780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTTACATCATCAATGCACATATTCTCAACA 839
|||
QY 241 GATCACGGTCACTAGTACACAGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 300
|||
Db 840 GATCACGGTCACTAGTACACAGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 899
|||
QY 301 GAGCTATGCCCTACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACACGGAGCCCAT 360
|||
Db 900 GAGCTATGCCCTACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACACGGAGCCCAT 959
|||
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 420
|||
Db 960 TCCTTCACAGCATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 1019
|||
QY 421 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 480
|||
Db 1020 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1079
|||
QY 481 TCCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
|||
Db 1080 TCCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
|||

QY 541 CCAGTTTCATACTCATGAGGGGTACATGATGATTTTACAGCCCATCAGGCGCGGTTGG 600
Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGATTTTACAGCCCATCAGGCGCGGTTGG 1199
QY 601 TAAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1200 TAAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 661 AACTGAAGTACAAGACGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 1260 AACTGAAGTACAAGACGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 721 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 1320 TAGCATGGAACAAACAAAGCAATTTACATAGA----- 1350
QY 781 GAAAAGATTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 840
Db 1351 ----- 1350
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 900
Db 1351 ----- 1350
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCAATGGTGGTAGT 960
Db 1351 ----- 1350
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 1351 ----- 1350
QY 1021 AGATCGATGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAC 1080
Db 1351 -----AC 1352
QY 1081 TCATFAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGA 1140
Db 1353 TCATFAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGA 1412
QY 1141 AGCTGAAACAACTGCCAATGTCTCTACAGGATGTAACCCGTAAAGGAAGGCTCCTAGAGA 1200
Db 1413 AGCTGAAACAACTGCCAATGTCTCTACAGGATGTAACCCGTAAAGGAAGGCTCCTAGAGA 1472
QY 1201 CTCGAAGGGAGTAAAGAGTGTATGAAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 1260
Db 1473 CTCGAAGGGAGTAAAGAGTGTATGAAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 1532
QY 1261 TCACACAGATGTTTATCACAACCTGGATGAAACAGCAAGGCTCCTGAGATCCCTGGA 1320
Db 1533 TCACACAGATGTTTATCACAACCTGGATGAAACAGCAAGGCTCCTGAGATCCCTGGA 1592
QY 1321 AGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
Db 1593 AGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1652
QY 1381 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCTCATTTGGAAGCCAGTTCTGACCAGTG 1440
Db 1653 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCTCATTTGGAAGCCAGTTCTGACCAGTG 1712
QY 1441 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGCTGGCTACAGCTGAAAGATGATGA 1500
Db 1713 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGCTGGCTACAGCTGAAAGATGATGA 1772
QY 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1560
Db 1773 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1832
QY 1561 ACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 1620
Db 1833 ACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 1892

QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAAACTCTACCAGGA 1680
Db 1893 GACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAAGGACTAGAGAAACTCTACCAGGA 1952
QY 1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
Db 1953 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 2012
QY 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACTCCGCTGACTGGCAGAG 1800
Db 2013 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACTCCGCTGACTGGCAGAG 2072
QY 1801 AAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGA 1860
Db 2073 AAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGA 2132
QY 1861 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1920
Db 2133 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 2192
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTCGCC 1980
Db 2193 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTCGCC 2252
QY 1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACCTTTGGGCAT 2040
Db 2253 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACCTTTGGGCAT 2312
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
Db 2313 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2372
QY 2101 GCAGGTGCGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 2160
Db 2373 GCAGGTGCGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 2432
QY 2161 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2220
Db 2433 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2492
QY 2221 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAAAACAACTTGCTGGGACCATCCCAA 2280
Db 2493 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAAAACAACTTGCTGGGACCATCCCAA 2552
QY 2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGATCTCTGAGCTTATAG 2340
Db 2553 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGATCTCTGAGCTTATAG 2612
QY 2341 GACTGCCATGAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db 2613 GACTGCCATGAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2672
QY 2401 A 2401
Db 2673 A 2673

RESULT 12
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX
AC AAD37258;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.

OS Chimeric.
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 62-63; 71pp; English.
XX
XC The present invention relates to an isolated nucleotide sequence encoding
XC a dystrophin minigene. The minigene comprises N-terminal or modified N-
XC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XC domains and cysteine-rich domains of dystrophin or utrophin genes. The
XC invention also relates to a recombinant adeno-associated virus (AAV)
XC comprising dystrophin minigene operably linked to an expression control
XC element. The dystrophin minigene in operable linkage with an expression
XC control element, in a recombinant adeno-associated virus or retrovirus is
XC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XC vector plasmid construct containing human dystrophin minigenes, a muscle
XC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;
Query Match 72.3%; Score 1737; DB 6; Length 4498;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2074; Conservative 0; Mismatches 0; Indels 327; Gaps 1;
2Y 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
DB 1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1416
2Y 61 TGTTGATACCACTTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
DB 1417 TGTTGATACCACTTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476
2Y 121 AGTTTTCCTCAACAAGTGAGCATTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
DB 1477 AGTTTTCCTCAACAAGTGAGCATTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536
2Y 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 240
DB 1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 1596
2Y 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 300
DB 1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 1656
2Y 301 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTTGACCCCTACACGGAGCCCAT 360
DB 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTTGACCCCTACACGGAGCCCAT 1716
2Y 361 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 420
DB 1717 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1776
2Y 421 TGAAGTAACTGGACCGGTTATCAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 480
DB 1777 TGAAGTAACTGGACCGGTTATCAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 1836

QY 481 TGCTGAGGACACATTCGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
DB 1837 TGCTGAGGACACATTCGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 541 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 600
DB 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA 660
DB 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA 2016
QY 661 AACTGAAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
DB 2017 AACTGAAAGTACAAGACAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 721 TAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
DB 2077 TAGCATGGAAAAACAAGCAATTTTACATAGA----- 2107
QY 781 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAAGAAAGAAATGGAGGAAGA 840
DB 2108 ----- 2107
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAAGTGCTTCA 900
DB 2108 ----- 2107
QY 901 AGAAGATCTAGAACACAGAACAGTCAAGGTCAATTTCTCTCATTACATGGTGGTGTAGT 960
DB 2108 ----- 2107
QY 961 TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
DB 2108 ----- 2107
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAC 1080
DB 2108 -----AC 2109
QY 1081 TCATAGATTACTGTCAACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1140
DB 2110 TCATAGATTACTGTCAACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 2169
QY 1141 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAGGAAAAAGGCTCCTAGAAGA 1200
DB 2170 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAGGAAAAAGGCTCCTAGAAGA 2229
QY 1201 CTCCAAGGGAGTAAAGAGCTGTATGAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 1260
DB 2230 CTCCAAGGGAGTAAAGAGCTGTATGAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 2289
QY 1261 TCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
DB 2290 TCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2349
QY 1321 AGGTTCCGATGATGCAGTCTCTCAACATTAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
DB 2350 AGGTTCCGATGATGCAGTCTCTCAACATTAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2409
QY 1381 TGAACCTTCGAAAAAAGTCTCTCAACATTAAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 1440
DB 2410 TGAACCTTCGAAAAAAGTCTCTCAACATTAAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 2469
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
DB 2470 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2529
QY 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGAAACGATGT 1560
DB 2530 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGAAACGATGT 2589

QY 1561 ACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGACTCTTGA 1620
Db |||||||
QY 1621 GACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGGA 1680
Db |||||||
QY 1681 GCTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGGA 1740
Db |||||||
QY 1741 GCTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGGA 1800
Db |||||||
QY 1801 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1860
Db |||||||
QY 1861 CCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT 1920
Db |||||||
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCC 1980
Db |||||||
QY 1981 TCTGAAAGAGAACTGAGCCACGTCATGACCTTGCTGCGCAGCTTACCACITTTGGGCAT 2040
Db |||||||
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCT 2100
Db |||||||
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACTTTGGTCC 2160
Db |||||||
QY 2161 AGCATCTCAGCACTTTCTTCCACGCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCC 2220
Db |||||||
QY 2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAACAACTTGCTGGGACCATCCCAA 2280
Db |||||||
QY 2281 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTCTAGCTTCTAGCTTATAG 2340
Db |||||||
QY 2341 GACTGCCATGAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db |||||||
QY 2401 A 2401
Db 3430 A 3430

RESULT 13
ABK81998
ID ABK81998 standard; DNA; 5339 BP.

XX ABK81998;
AC
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS Example 6; Fig 13; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention
XX
SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;

Query Match 71.8%; Score 1725; DB 6; Length 5339;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2068; Conservative 0; Mismatches 0; Indels 333; Gaps 1;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db |||||||
QY 799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 858
Db |||||||
QY 61 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
Db |||||||
QY 859 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 918
Db |||||||
QY 121 AGTTTGGCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db |||||||
QY 919 AGTTTGGCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 978
Db |||||||
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA 240
Db |||||||
QY 979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA 1038
Db |||||||
QY 241 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
Db |||||||
QY 1039 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1098
Db |||||||
QY 301 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACCTCTGACCCTACACGGAGCCCAT 360
Db |||||||
QY 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACCTCTGACCCTACACGGAGCCCAT 1158
Db |||||||
QY 361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 420
Db |||||||
QY 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
Db |||||||
QY 421 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC 480
Db |||||||
QY 1219 TGAAGTAAACCTGGACCGTTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278
Db |||||||

XX OS Homo sapiens.
XX WO200183695-A2.
XX PN
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US013677.
XX PR 28-APR-2000; 2000US-0200777P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX DR WPI; 2002-049342/06.
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 51-52; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is human
XX dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
XX hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
XX CR domain) and 11047-11058 (dystrophin last 3 amino acids)
SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
Query Match 70.6%; Score 1695; DB 6; Length 3510;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
DB 600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 659
QY 61 TGTGTATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCATCTCTTCCA 120
DB 660 TGTGTATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCATCTCTTCCA 719
QY 121 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
DB 720 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 240
DB 780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 839
QY 241 GATCAGGTCAGTCTAGCAGGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 300
DB 840 GATCAGGTCAGTCTAGCAGGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 899
QY 301 GAGTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 360
DB 900 GAGTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 959
QY 361 TCCTTACAGCATTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 420
DB 960 TCCTTACAGCATTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1019
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480

DB 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
QY 481 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 540
DB 1080 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
QY 541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 600
DB 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1199
QY 601 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 660
DB 1200 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1259
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
DB 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 721 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
DB 1320 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAAAGAAACAAGGAAAAATGGAGGAAGA 840
DB 1380 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAAAGAAACAAGGAAAAATGGAGGAAGA 1439
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAAGGTGCTTCA 900
DB 1440 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAAGGTGCTTCA 1499
QY 901 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 960
DB 1500 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 1559
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
DB 1560 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGACAC 1080
DB 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTCTTTTACAAGA--- 1676
QY 1081 TCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1140
DB 1677 ----- 1676
QY 1141 AGCTGAACAACACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
DB 1677 ----- 1676
QY 1201 CTCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1260
DB 1677 ----- 1676
QY 1261 TCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
DB 1677 ----- 1676
QY 1321 AGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1380
DB 1677 ----- 1676
QY 1381 TGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
DB 1677 -----CAGTTCTGACCAGTG 1691
QY 1441 GAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGA 1500
DB 1692 GAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGA 1751
QY 1501 ATTAAGCCGCGCAGGCACCTATTGGAGCGGACCTTTCAGCAGCTTCAGAAAGCAGAACGATGT 1560

Db 1752 ATTAAGCCGGCAGGCACCTATTGGAGGGGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1811
Qy 1561 ACATAGGGCCCTTCAAGAGGGAAATTGAAAACCTGTAATCATGAGTACTCTTGA 1620
Db 1812 ACATAGGGCCCTTCAAGAGGGAAATTGAAAACCTGTAATCATGAGTACTCTTGA 1871
Qy 1621 GACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
Db 1872 GACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1931
Qy 1681 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1740
Db 1932 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1991
Qy 1741 GGCTGAGGAGGTCAATACTAGTGGGAAATAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
Db 1992 GGCTGAGGAGGTCAATACTAGTGGGAAATAATTGAACCTGCACTCCGCTGACTGGCAGAG 2051
Qy 1801 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGTGGA 1860
Db 2052 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGTGGA 2111
Qy 1861 CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCTTGGAGCCCCGTGGGCGATCTCCT 1920
Db 2112 CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCTTGGAGCCCCGTGGGCGATCTCCT 2171
Qy 1921 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
Db 2172 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 2231
Qy 1981 TCTGAAAGAGAGAGCTGAGCCACGTCATGACCTTGTCTGCGAGCTTACCACCTTTGGGCAT 2040
Db 2232 TCTGAAAGAGAGAGCTGAGCCACGTCATGACCTTGTCTGCGAGCTTACCACCTTTGGGCAT 2291
Qy 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 2100
Db 2292 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 2351
Qy 2101 GCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 2160
Db 2352 GCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 2411
Qy 2161 AGCATCTCAGCACTTTCTTTCCACGCTGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCC 2220
Db 2412 AGCATCTCAGCACTTTCTTTCCACGCTGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCC 2471
Qy 2221 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCA 2280
Db 2472 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCA 2531
Qy 2281 AATGACAGAGTCTTACCAGTCTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAG 2340
Db 2532 AATGACAGAGTCTTACCAGTCTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAG 2591
Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTC 2400
Db 2592 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTC 2651
Qy 2401 A 2401
Db 2652 A 2652

RESULT 15
AAD37259 standard; DNA; 4476 BP.
AAD37259;
21-AUG-2002 (first entry)
Adeno-associated virus vector plasmid, AAV-MCK-3510.

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 63-65; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (CK) promoter and a small polyA signal sequence
XX
SQ Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

Query Match 70.6%; Score 1695; DB 6; Length 4476;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

Qy 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 1356 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1415
Qy 61 TGTGTATACCACCTATCCAGATAAGAAGTCCATCTTAATGTATCATCATCATCTCTTCCA 120
Db 1416 TGTGTATACCACCTATCCAGATAAGAAGTCCATCTTAATGTATCATCATCATCTCTTCCA 1475
Qy 121 AGTTTTCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 1476 AGTTTTCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1535
Qy 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACATATTCTCAACA 240
Db 1536 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACATATTCTCAACA 1595
Qy 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCCTCGATTCAA 300
Db 1596 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCCTCGATTCAA 1655
Qy 301 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCA 360
Db 1656 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCA 1715
Qy 361 TCCTTCACAGCAATTTGGAAGCTCCTCGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 420

Db 1716 TCCCTTACAGCAATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1775
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCCTGGCTTCTTTTC 480
Db 1776 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCCTGGCTTCTTTTC 1835
QY 481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAAGTGGTGAAGA 540
Db 1836 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAAGTGGTGAAGA 1895
QY 541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 600
Db 1896 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 1955
QY 601 TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAAGATGAAGA 660
Db 1956 TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAAGATGAAGA 2015
QY 661 AACTGGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 720
Db 2016 AACTGGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 2075
QY 721 TAGCATGGAAAAACAAAAGCAATTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAAC 780
Db 2076 TAGCATGGAAAAACAAAAGCAATTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAAC 2135
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 840
Db 2136 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 2195
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAACATAAAGGTGCTTCA 900
Db 2196 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAACATAAAGGTGCTTCA 2255
QY 901 AGAAGATCTAGAACAAGAACAACTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 960
Db 2256 AGAAGATCTAGAACAAGAACAACTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 2315
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACITTAAGGTATTGGG 1020
Db 2316 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACITTAAGGTATTGGG 2375
QY 1021 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
Db 2376 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA --- 2432
QY 1081 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGAAAGTTTCTTGGCTGGCTTACAGA 1140
Db 2433 ----- 2432
QY 1141 AGCTGAACAACACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGTCCTAGAAGA 1200
Db 2433 ----- 2432
QY 1201 CTCCAAGGAGTAAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATGAAGC 1260
Db 2433 ----- 2432
QY 1261 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCCTGAGATCCCTGGA 1320
Db 2433 ----- 2432
QY 1321 AGGTTCCGATGATGCAGTCTGTTCAAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1380
Db 2433 ----- 2432
QY 1381 TGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAGTG 1440
Db 2433 -----CAGTTCTGACCAGTG 2447
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGA 1500

Db 2448 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAAGATGATGA 2507
QY 1501 ATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTTCAGCAGTTCAGAAAGCAGAACGATGT 1560
Db 2508 ATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTTCAGCAGTTCAGAAAGCAGAACGATGT 2567
QY 1561 ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
Db 2568 ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2627
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACTCTACCAAGGA 1680
Db 2628 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACTCTACCAAGGA 2687
QY 1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
Db 2688 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2747
QY 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1800
Db 2748 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 2807
QY 1801 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGA 1860
Db 2808 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGA 2867
QY 1861 CCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCT 1920
Db 2868 CCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCT 2927
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCC 1980
Db 2928 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCC 2987
QY 1991 TCTGAAAGAGAACGTCGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2040
Db 2998 TCTGAAAGAGAACGTCGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 3047
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 2100
Db 3048 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 3107
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCCACAGGGACTTTGGTCC 2160
Db 3108 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCCACAGGGACTTTGGTCC 3167
QY 2161 AGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCCTGGGAGAGGCCATCTCGCC 2220
Db 3168 AGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCCTGGGAGAGGCCATCTCGCC 3227
QY 2221 AAACAAAAGTGCCTACTATATCAACCCAGAGACTCAAAACAACTTGTCTGGGACCATCCCAA 2280
Db 3228 AAACAAAAGTGCCTACTATATCAACCCAGAGACTCAAAACAACTTGTCTGGGACCATCCCAA 3287
QY 2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAAATGATGTGATTTCTCAGCTTATAG 2340
Db 3288 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAAATGATGTGATTTCTCAGCTTATAG 3347
QY 2341 GACTGCCATGAAAACTCCGAAGACTGCAGAAGGCCCTTTTGGTGGATCTCTTGAGCCTGTC 2400
Db 3348 GACTGCCATGAAAACTCCGAAGACTGCAGAAGGCCCTTTTGGTGGATCTCTTGAGCCTGTC 3407
QY 2401 A 2401
Db 3408 A 3408

Search completed: April 4, 2004, 18:20:16
Job time : 656.552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:58:13 ; Search time 116.44 Seconds
(without alignments)
11443.127 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000

Perfect score: 2401

Sequence: 1 acatgcattcaacatcgcca.....tggatctcttgagcctgtca 2401

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match %	Length	DB ID	Description
1	1335.6	55.6	5952	4	US-09-687-875A-1
2	1335.6	55.6	13977	4	US-09-484-970B-60
3	1143.6	47.6	19307	3	US-08-836-022A-10
4	1143.6	47.6	19307	3	US-09-427-048A-10
5	487	20.3	6045	4	US-09-091-501B-7
6	487	20.3	10320	4	US-09-091-501B-9
7	401	16.7	3915	4	US-09-976-594-93
8	79.4	3.3	200	4	US-09-091-501B-5
9	78.6	3.3	200	4	US-09-091-501B-4
10	78.6	3.3	200	4	US-09-091-501B-6
11	76.6	3.2	7218	1	US-08-232-463-14
12	63.6	2.6	238	4	US-09-687-875A-13
13	46.2	1.9	505	4	US-09-621-976-15639
14	44.2	1.8	2574	4	US-09-668-313A-10
15	44	1.8	1230025	4	US-09-198-452A-1
16	43.4	1.8	1179	4	US-09-107-532A-1186
17	42.8	1.8	1690	4	US-09-620-312D-69
18	42.8	1.8	7812	3	US-09-368-590-1
19	40.4	1.7	832	4	US-09-621-976-2813
20	40.4	1.7	2223	1	US-08-257-073-4
21	39.6	1.6	6960	2	US-08-841-349-3
22	39.6	1.6	6960	4	US-09-431-184A-3
23	39.6	1.6	8176	2	US-08-841-349-5
24	39.6	1.6	8176	4	US-09-431-184A-5
25	39.2	1.6	16995	4	US-08-961-527-82
26	38.6	1.6	1751	4	US-09-620-312D-847
27	38.6	1.6	1995	1	US-08-425-069-3

28	38.6	1.6	1995	2	US-08-317-844B-3	Sequence 3, Appli
29	38.4	1.6	7672	4	US-09-220-132-24	Sequence 24, Appl
30	38.2	1.6	428	4	US-09-668-313A-3	Sequence 3, Appli
31	38.2	1.6	1131	6	5180810-3	Patent No. 5180810
32	38.2	1.6	1784	6	5180810-2	Patent No. 5180810
33	38.2	1.6	4439	4	US-09-668-313A-17	Sequence 17, Appl
34	38	1.6	1394	4	US-09-247-155-76	Sequence 76, Appl
35	36.8	1.5	1886	6	5210183-1	Patent No. 5210183
36	36.6	1.5	1845	4	US-08-887-534A-22	Sequence 22, Appl
37	36.6	1.5	1845	4	US-09-527-431-22	Sequence 22, Appl
38	36.6	1.5	7075	4	US-08-956-171E-263	Sequence 263, App
39	36.2	1.5	289	3	US-09-007-005-17	Sequence 17, Appl
40	36.2	1.5	289	3	US-09-244-796-17	Sequence 17, Appl
41	36.2	1.5	713	4	US-09-091-501B-2	Sequence 2, Appli
42	36.2	1.5	2447	2	US-09-014-969-14	Sequence 14, Appl
43	36.2	1.5	11049	4	US-10-204-708-23	Sequence 23, Appl
44	36.2	1.5	168575	4	US-09-426-290-1	Sequence 1, Appli
45	36	1.5	724	4	US-08-956-171E-832	Sequence 832, App

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Paul
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
; US-09-687-875A-1

Query Match 55.6%; Score 1335.6; DB 4; Length 5952;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	1028	TGGGCAACATCTGTAGATGGACAGACACCGCTGGGTTCTTTTACAAGACACTCATAGA	1087
Db	2902	TGGAGAAGCATTCTATAAAGGGTGAGTGACGAGAGGCTGCTTTGGAAGAACTCATAGA	2961
Qy	1088	TTACTGCAACAGTTCCTCCCTGGACCTGGACCTGGAAGTTTCTTGCTGGCTTACAGAAGCTGAA	1147
Db	2962	TTACTGCAACAGTTCCTCCCTGGACCTGGACCTGGAAGTTTCTTGCTGGCTTACAGAAGCTGAA	3021
Qy	1148	ACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGACTCCAAG	1207
Db	3022	ACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGACTCCAAG	3081
Qy	1208	GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA	1267
Db	3082	GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA	3141
Qy	1268	GATGTTTATCACAACCTGGATGAAACACAGCAAAAATCCTGAGATCCCTGGAAGTTCC	1327

Db 3142 GATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 3201

Qy 1328 GATGATGCAGTCTCTGTACAAAGACGTTTGGTAACATGAACCTTCAAGTGGAGTGAACCTT 1387

Db 3202 GATGATGCAGTCTCTGTACAAAGACGTTTGGTAACATGAACCTTCAAGTGGAGTGAACCTT 3261

Qy 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 1447

Db 3262 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 3321

Qy 1448 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAAAGATGATGAATTAAGC 1507

Db 3322 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAAAGATGATGAATTAAGC 3381

Qy 1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGCTTTCAGAAAGCAGACGATGTACATAGG 1567

Db 3382 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGCTTTCAGAAAGCAGACGATGTACATAGG 3441

Qy 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1627

Db 3442 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 3501

Qy 1628 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTTACCAAGAGCCCA 1687

Db 3502 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTTACCAAGAGCCCA 3561

Qy 1688 GAGCTGCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747

Db 3562 GAGCTGCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 3621

Qy 1748 GAGTCAATACTGAGTGGGAAAAATTGAACCTCGCTGACTCGGCTGAGTGGCAGAGAAAAATA 1807

Db 3622 GAGTCAATACTGAGTGGGAAAAATTGAACCTCGCTGACTCGGCTGAGTGGCAGAGAAAAATA 3681

Qy 1808 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG 1867

Db 3682 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG 3741

Qy 1868 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGGCGATCTCCTCATTTGAC 1927

Db 3742 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGGCGATCTCCTCATTTGAC 3801

Qy 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAA 1987

Db 3802 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAA 3861

Qy 1988 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 2047

Db 3862 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 3921

Qy 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 2107

Db 3922 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 3981

Qy 2108 GCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 2167

Db 3982 GCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 4041

Qy 2168 CAGCACCTTTCTTTCCAGCTGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227

Db 4042 CAGCACCTTTCTTTCCAGCTGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 4101

Qy 2228 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACA 2287

Db 4102 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACA 4161

Qy 2288 GAGCTCTACAGTCTTTAGTGAACCTGAATATGTGAGATTTCTCAGCTTATAGGACTGCC 2347

Db 4162 GAGCTCTACAGTCTTTAGTGAACCTGAATATGTGAGATTTCTCAGCTTATAGGACTGCC 4221

Qy 2348 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCA 2401

Db 4222 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCA 4275

RESULT 2

US-09-484-970B-60

; Sequence 60, Application US/09484970B

; Patent No. 6426186

; GENERAL INFORMATION:

; APPLICANT: Jones, Karen A.

; APPLICANT: Volkmuth, Wayne

; APPLICANT: Walker, Michael G.

; TITLE OF INVENTION: BONE REMODELING GENES

; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B

; CURRENT FILING DATE: 2000-01-18

; NUMBER OF SEQ ID NOS: 172

; SOFTWARE: PERL Program

; SEQ ID NO 60

; LENGTH: 13977

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1

; NAME/KEY: unsure

; LOCATION: 11721-11761, 12294, 13969

; OTHER INFORMATION: a, t, c, g, or other

US-09-484-970B-60

Query Match 55.6%; Score 1335.6; DB 4; Length 13977;

Best Local Similarity 98.3%; Pred. No. 0;

Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1028 TGGGCAACATCTGTAGATGACAGAGACCGCTGGTCTCTTTTACAGACACTCATAGA 1087

Db 8216 TGGAGAGCATTCATAAAGGGTGAGTGAGCGAGAGGCTGCTTTGGAAAGAACTCATAGA 8275

Qy 1088 TTACTGCAACAGTTCCCTCGACCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 1147

Db 8276 TTACTGCAACAGTTCCCTCGACCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 8335

Qy 1148 ACACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCCAAG 1207

Db 8336 ACACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCCAAG 8395

Qy 1208 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 1267

Db 8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 8455

Qy 1268 GATGTTTATCACAACTGTGATGAAACACAGCCAAAATCCTGAGATCCCTGGAAGGTTCC 1327

Db 8456 GATGTTTATCACAACTGTGATGAAACACAGCCAAAATCCTGAGATCCCTGGAAGGTTCC 8515

Qy 1328 GATGATGCAGTCTCTTCAACATAGGTCCCATTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 1387

Db 8516 GATGATGCAGTCTCTTCAACATAGGTCCCATTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 8575

Qy 1388 CGGAAAAAGTCTCTCAACATAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT 1447

Db 8576 CGGAAAAAGTCTCTCAACATAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT 8635

Qy 1448 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507

Db 8636 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695

Qy 1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGCTTTCAGAAAGCAGAACGATGTACATAGG 1567

Db 8696 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGCTTTCAGAAAGCAGAACGATGTACATAGG 8755

Qy 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGAGACTGTA 1627

Db 8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGAGACTGTA 8815

Qy 1628 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTTACCAGGAGCCCA 1687

||||| 8816 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGAGTCTAGAGAACTCTACCGAGGAGCCAGAG 8875
||||| 1688 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
||||| 8876 GAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
||||| 1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA 1807
||||| 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA 8995
||||| 1808 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
||||| 8996 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055
||||| 1868 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 1927
||||| 9056 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 9115
||||| 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAA 1987
||||| 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAA 9175
||||| 1988 GAGAACGTGAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC 2047
||||| 9176 GAGAACGTGAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC 9235
||||| 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGAGTG 2107
||||| 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGAGTG 9295
||||| 2108 GCCGTGAGGAGCCAGTCAGGCAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCT 2167
||||| 9296 GCCGTGAGGAGCCAGTCAGGCAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCT 9355
||||| 2168 CAGCACTTTCTTCCAGCTGTGTCCAGGTCCTGGGAGAGGCCATCTCGCCAAACAAA 2227
||||| 9356 CAGCACTTTCTTCCAGCTGTGTCCAGGTCCTGGGAGAGGCCATCTCGCCAAACAAA 9415
||||| 2228 GTGCCCTACTATATCAACCAGAGACTCAAAACAACTTGTGGGACCATCCCAAATGACA 2287
||||| 9416 GTGCCCTACTATATCAACCAGAGACTCAAAACAACTTGTGGGACCATCCCAAATGACA 9475
||||| 2288 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGAGATTTAGCTTATAGGACTGCC 2347
||||| 9476 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGAGATTTAGCTTATAGGACTGCC 9535
||||| 2348 ATGAAACTCCGAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 2401
||||| 9536 ATGAAACTCCGAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 9589

RESULT 3
JS-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-836-022A-10
Query Match 47.6%; Score 1143.6; DB 3; Length 19307;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 1230; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 1028 TGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
Db 6478 TGGGAAAACATTCATAAAGAGTAAGTGAGCAAGAGGCTGCTTTGGAAAGAACTCATAGA 6419
QY 1088 TTAAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 1147
Db 6418 TTAAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 6359
QY 1148 ACAAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 1207
Db 6358 ACAAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 6299
QY 1208 GAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 1267
Db 6298 GAGTCAACAGTTCCCGCTGGACCTGGAAAGTTTCTTGGCTTACAGAAAGCTGAA 6239
QY 1268 GATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 1327
Db 6238 GATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 6179
QY 1328 GATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 1387
Db 6178 GATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 6119
QY 1388 CGGAAAAAGTCTCTCAACATTTAGGTCCTTGGAGCCAGTTCTGACCCAGTGGAGCCGT 1447
Db 6118 CGGAAAAAGTCTCTCAACATTTAGGTCCTTGGAGCCAGTTCTGACCCAGTGGAGCCGT 6059
QY 1448 CTGCACCTTTCTCTGCAGGAACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
Db 6058 CTGCACCTTTCTCTGCAGGAACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 5999
QY 1508 CGGAGGCACCTATTGGAGCCGACTTTCCAGCAGTTTCCAGAGCAAGATGATGATGATGATG 1567
Db 5998 CGGAGGCACCTATTGGAGCCGACTTTCCAGCAGTTTCCAGAGCAAGATGATGATGATGATG 5939
QY 1568 GCCTTCAAGAGGGAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA 1627
Db 5938 GCCTTCAAGAGGGAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA 5879
QY 1628 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTTACAGAGAGCCAGAG 1687
Db 5878 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTTACAGAGAGCCAGAG 5819

QY 1688 GAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
|||
Db 5818 GAACTGCCTCCTGAAGAAAGAGCTCAGAAATGTCACTCGGCTCCTACGAAAGCAGGCTGAA 5759
|||
QY 1748 GAGGTCAATACTGAGTGGGAAATAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATA 1807
|||
Db 5758 GAGGTCAACGCTGAATGGGACAAATTTGAACCTGCCTCAGCTGATGGCAGAGAAATA 5699
|||
QY 1808 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
|||
Db 5698 GATGAAGCTCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 5639
|||
QY 1868 CTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGATCTCCTCAATTGAC 1927
|||
Db 5638 TTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCAGTGGGGATCTCCTCAATTGAC 5579
|||
QY 1928 TCTCTCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAA 1987
|||
Db 5578 TCTCTCAAAGATCACCTTGAAAAAGTCAAGGCACTTCGGGAGAAATTTGACCTCTTAAA 5519
|||
QY 1988 GAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTC 2047
|||
Db 5518 GAGAATGTCAATCGTGTCAATGACCTTGCACATCAGCTGACCACTCTGGGCAATTCAGCTC 5459
|||
QY 2048 TCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGCAGGTG 2107
|||
Db 5458 TCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAAGATGGAGCTTCTACAGGTG 5399
|||
QY 2108 GCCGTGAGGACCGAGTCAAGGAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 2167
|||
Db 5398 GCTGTGAGGACCGTGTCAAGCAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCATCC 5339
|||
QY 2168 CAGCACTTCTTTCCACGCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227
|||
Db 5338 CAGCACTTCTTTCCACCTTCACTTCAAGGCTCCCTGGGAGAGAGCCATCTCACCAACAAA 5279
|||
QY 2228 GTGCCCTACTATATCAACCAAGAGACTCAACCAACTTGTCTGGGACCACTCCCAAAATGACA 2287
|||
Db 5278 GTGCCCTACTATATCAACCAAGAGACTCAACCAACTTGTCTGGGACCACTCCCAAAATGACA 5219
|||
QY 2288 GAGCTTACCAGTCTTTTAGTGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCC 2347
|||
Db 5218 GAGCTTACCAGTCTTTTAGTGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCC 5159
|||
QY 2348 ATGAAACTCCGAAGACTGCAAGAGCCCTTTTGGTGGATCTCTTGGCTGTCA 2401
|||
Db 5158 ATGAACTCAGAAAGGCTCAGAAAGGCCCTTTTGGTGGATCTCTTGGCTGTCA 5105
|||

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975

GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania

Wilson, James M.

Fisher, Krishna J.

Chen, Shu-Jen

Weitzman, Matthew

TITLE OF INVENTION: Improved Adenovirus Virus and

Methods of Use Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P O Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 47.6%; Score 1143.6; DB 3; Length 19307;

Best Local Similarity 89.5%; Pred. No. 0;

Matches 1230; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1028 TGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
|||
Db 6478 TGGGCAAAACATCTATAAAGAGTAAGTGAGCAAGAGGCTGCTTTGGAAGAACTCATAGA 6419
|||
QY 1088 TTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGGCTTGGCTTACAGAAGCTGAA 1147
|||
Db 6418 TTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTCTGGATTACGAAGCAGAA 6359
|||
QY 1148 ACAACTGCCAATGTCTACAGATGCTACCCGTAGGAAAGGCTCCTAGAGACTCCAAG 1207
|||
Db 6358 ACAACTGCCAATGTCTACAGAGCTTCCCGTAGGAGAGCTCCTAGAGACTCCAAG 6299
|||
QY 1208 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 1267
|||
Db 6298 GGAGTCAGAGAGCTGATGAACAATGGCAAGATCTCCAAGGAGAAATTTGAACCTCACACA 6239
|||
QY 1268 GATGTTTATCAACCTGGATGAAAACAGCCAAATAATCCTGAGATCCCTGGAAGGTTCC 1327
|||
Db 6238 GATATCTATCAAACTCTTGATGAAAATGGCCAAAATAATCCTGAGATCCCTGGAAGGTTCCG 6179
|||
QY 1328 GATGATGAGTCTCTGTACAAAGACGTTTGGATACATGAACCTTCAAGTGAGTGAACCTT 1387
|||
Db 6178 GATGAAGCAGCCCTGTTTACAAGACGTTTGGATACATGAATTTCAAGTGAGTGAACCTT 6119
|||
QY 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTTCTGACCAGTGAAGCGT 1447
|||
Db 6118 CAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTTCTGACCAGTGAAGCGT 6059
|||
QY 1448 CTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
|||
Db 6058 TTGCATCTTCTCTTCAAGAACTTCTTGTGGCTACAGCTGAAAGATGATGAAGTGAAGC 5999
|||
QY 1508 CGGCAGGCACTTATTTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGATACATAGG 1567
|||
Db 5998 CGTCAGGCACTTATTTGGAGGCGACTTTCAGCAGTTCAGAAAGCAGAACGATGATACATAGG 5939
|||
QY 1568 GCCTTCAAGAGGGAATTTGAAAATAAAGAACCTGTAAATCATGAGTACTCTGAGACTGTA 1627
|||
Db 5938 GCCTTCAAGAGGGAATTTGAAAATAAAGAACCTGTAAATCATGAGTACTCTGAGACTGTG 5879
|||
QY 1628 CGAATATTTCTGACAGAGCAGGCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCGAGA 1687
|||
Db 5878 AGAATATTTCTGACAGAGCAGGCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCGAGA 5819
|||

QY 1688 GAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
|||
Db 5818 GAACTGCCCTCTGAAGAAAGAGCTCAGAATGTCACTCGGCTCTCTACGAAAGCAGGCTGAA 5759
|||
QY 1748 GAGGTCAATCTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA 1807
|||
Db 5758 GAGGTCAACGCTGAATGGACAAATTGAACCTCGCTCAGCTGATTGGCAGAGAAAAATA 5699
|||
QY 1808 GATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGATGAGCTGGACCTCAAG 1867
|||
Db 5698 GATGAAGCTCTTGAAGAGACTCCAGGAACCTCAGGAAGCTGCCGATGAACCTGGACCTCAAG 5639
|||
QY 1868 CTGCGCCAAAGTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCATTTGAC 1927
|||
Db 5638 TTGCGCCAAAGTGAGGTGATCAAGGGATCTTGGCAGCCAGTGGGGATCTCTCATTTGAC 5579
|||
QY 1928 TCTCTCCAAAGTCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAA 1987
|||
Db 5578 TCTCTGCAAGATCACCTTGAAAAAGTCAAGGCACCTTCGAGGAGAAATTGCACTCTTAAA 5519
|||
QY 1988 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC 2047
|||
Db 5518 GAGAAATGTAATCGTGTCAATGACCTTGACATCAGCTGACCACTGGGCATTCAGCTC 5459
|||
QY 2048 TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACAGATGGAAGCTTTCAGGCTG 2107
|||
Db 5458 TCACCTTATAACCTCAGCACCTTTGGAAGATCTGAATACCAGATGGAGGCTTCTACAGGTG 5399
|||
QY 2108 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT 2167
|||
Db 5398 GCTGTGAGGACCGTGTGACACAGCTGCATGAAGCCACAGGGACTTTGGTCCATGCC 5339
|||
QY 2168 CAGCACTTTCTTTCCACGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2227
|||
Db 5338 CAGCACTTTCTTTCCACTTCAGTTTCAGGGTCCCTGGGAGAGAGCCATCTCACCACAAA 5279
|||
QY 2228 GTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 2287
|||
Db 5278 GTGCCCTACTATATCAACCCAGAGACCCAAACCACTTGTGGGACCAACCAAAATGACA 5219
|||
QY 2288 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCC 2347
|||
Db 5218 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGTTCCTCGCGTATAGGACTGCC 5159
|||
QY 2348 ATGAACCTCCGAAGACTGCGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 2401
|||
Db 5158 ATGAACCTCAGAAGGCTCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 5105
|||

RESULT 5
US-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Uterophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 20.3%; Score 487; DB 4; Length 6045;
Best Local Similarity 58.3%; Pred. No. 1.2e-139;
Matches 878; Conservative 0; Mismatches 620; Indels 9; Gaps 1;
QY 894 TGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGG 953
Db 2871 TACTGCTTCAAGAACTGGGTCTCTGGAGATGGTATCGTTCATGGCTTCGATAACGTCCTGC 2930
|||
QY 954 TGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAAACAACCTTAAGG 1013
Db 2931 AGAAACTCCTGGAGGAATATGGGAGTGATGACACAAGGAATGTGAAAGAAACCCACAGAGT 2990
|||
QY 1014 TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGTTCTTTTAC 1073
Db 2991 ACTTAAAAACATCATGGATCAATCTCAAAACAAGTATTGCTGACAGACAGAACGCCCTGG 3050
|||
QY 1074 AAGACACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGGCCTGGC 1133
Db 3051 AGGCTGAGTGGAGGACGGTGCAGGCCCTCTGCAGAGATCTGAAAAAATCTCTGAGAGTGA 3110
|||
QY 1134 TTACAGAAAGCTGAAAAAAGTCCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCTCC 1193
Db 3111 TCCAAGAAAGCAGAGACCCACAGTGAATGTGCTTGTGGATGCTCTCATCGGGAGATGCTC 3170
|||
QY 1194 TAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAAGTGAAA 1253
Db 3171 TTCAGGATAGTATCTTTGGCCAGGAACTCAACAGCAGATCGAGGACATCCAGGCAGAAA 3230
|||
QY 1254 TTGAAGCTCACACAGATGTTTTATCAACACCTGGATGAAAACAGCCCAAAAAATCCTGAGAT 1313
Db 3231 TTGATGCCCAATGACATATTTAAAGCATTTGACGGAACATTCAGGAAACAGGACAGATGTTAAAG 3290
|||
QY 1314 CCTTGAAGGTTCCGATGATGAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTG 1433
Db 3291 CTTTGGGAAATCTGAAGAGGCTACTATGCTTCAACATCGACTGGATGATGATGAACCAAA 3350
|||
QY 1374 AGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTG 1433
Db 3351 GATGGAATGACTTAAAGCAAAATCTGCTAGCATCAGGGCCCATTTGGAGGCCAGGCTG 3410
|||
QY 1434 ACCAGTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAG 1493
Db 3411 AGAAGTGGAAACAGGTTGCTGATGCTCTTAGAAGAACTGATCAAAATGGCTGAATATGAAAG 3470
|||
QY 1494 ATGATGAATTAAGCCGCGGACCTATTTGGAGGCGACTTTCCAGCAGTTCCAGAGCAGA 1553
Db 3471 ATGAAGAGCTTAAAGAAACAAATGCCTATTGGAGGAGATGTTCCAGCCTTACAGCTCCAGT 3530
|||
QY 1554 ACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTATCATGAGTA 1613
Db 3531 ATGACCATTTGTAAGGCCCTTGAGACGGGAGTTAAAGGAGAAAGATATTCTGTCTCTGAATG 3590
|||
QY 1614 CTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAC--- 1670
Db 3591 CTGTCGACCCAGGCCCGAGTTTCTTGGCTGATCAGCCCAATTTAGGCCCTTGAAGAGCCAA 3650
|||
QY 1671 -----TCTACCAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGATGTCATC 1724
Db 3651 GAAGAAACCTTACATCAAAAACAGAAATTAACCTCTGAGAGAGAGCCCAAAAGATTGCCA 3710
|||

1725 GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACT 1784
1786 AAGCCATGGCCAAACAGTCTTCTGAAGTCAAAGAAAAATGGGAAAGTCTAAATGCTGTAA 8045
1785 CCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGG 1844
8046 CTAGCAATTGGCAAAAGCAAGTGGACAAGGCATTGGAGAAACTCAGAGACCTGCAGGGAG 8105
1845 CCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGC 1904
8106 CTATGGATGACCTGGACGCTGACATGAAGGAGGCGAGTCCGTGCGGAATGGCTGGAAGC 8165
1905 CCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTC 1964
8166 CCGTGGGAGACTTACTCATTTGACTCGCTGCAGGATCACATTGAAAAAATCATGGCATTTA 8225
1965 GAGGAGAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCGAATGACCTTTCGTCGCCAGC 2024
8226 GAGAAGAAATTGCACCAATCAACTTTAAAGTTAAAACGGTGAATGATTTATCCAGTCAGC 8285
2025 TTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACA 2084
8286 TGTCTCCACTTGACCTGCTCATCCCTCTCTAAAGATGCTCGCCAGCTAGATGACCTTAATA 8345
2085 CCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCC 2144
8346 TGCATGGAACCTTTTACAGGTTTCTGTGGATGATCGCCCTTAAACAGCTTCAGGAAGCCC 8405
2145 ACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGTCCTTGGG 2204
8406 ACAGAGATTTTGGACCATCTCTCAGCATTTTCTCTACGTCAGTCCAGCTGCCGTGGC 8465
2205 AGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACAATT 2264
8466 AAAGATCCATTTCACATAATAAAGTGCCTATTACATCAACCATCAAAACACAGACCACCT 8525
2265 GCTGGGACCATCCCAAAATGACAGAGCTCTACCAGCTTTTAGCTGACCTGAATAATGTCA 2324
8526 GTTGGGACCATCTTAAATGACCCGAACCTCTTCAATCCCTTGTGCTGACCTGAATAATGTAC 8585
2325 GATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTGTCTGG 2384
8586 GTTTTCTGCTACCGTACAGCAATCAAAATCCGAAGACTCAAAAGCACTATGTTTGG 8645
2385 ATCTCTT 2391
8646 ATCTCTT 8652

RESULT 7
US-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 6673549 290344.1
US-09-976-594-93

Query Match 16.7%; Score 401; DB 4; Length 3915;
Best Local Similarity 62.5%; Pred. No. 3.8e-113;
Matches 644; Conservative 0; Mismatches 385; Indels 2; Gaps 1;
QY 1363 CATGAACTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGA 1422
Db 544 CATGAATCTGTGTGAATGAAATAAAAAAAGTCTCAACACTCCGCGCTCGCCTAGA 603
QY 1423 AGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTTGGTGTGGCT 1482
Db 604 GGCCTTCTCAGACCAACAGTGGAAAGCTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCT 663
QY 1483 ACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCAGGAGT 1542
Db 664 CAGCCAAAAGGATGAGGAGTTGTCAAGTCAAGTCCCTACAGGGGATGTGGCCCTGGT 723
QY 1543 TCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAAACCTGT 1602
Db 724 GCAACAGGAGAGGAGACACATCGCGCCTTTATGGAAGAAAGTCAAGTCTCGGGGCCCTTA 783
QY 1603 AATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACT 1662
Db 784 CATCTATTCTGTGTGGAGTCAGCTCAGGCTTCTGTGCCAGCACCTTGTGAGGAGT 843
QY 1663 AGAGAAACTCTACAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCAC 1722
Db 844 AGAGGAGCCTCAITCTGAGAGCAAGATACCTCCCCGAAACAGCGGATCCAGAAATCTCAG 903
QY 1723 TCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCA 1782
Db 904 CCGCTTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACCTGTGGGAGAAAGTTTGACAGCCCG 963
QY 1783 CTCCGCTGACTGSCAGAGAAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGA 1842
Db 964 CTGTGTGGACCAGCACCGTCACTTGTAGCGGACTCTGGAGCAGCTCTTGGAGATTCA--G 1021
QY 1843 GGCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGSCA 1902
Db 1022 GGCATGGAGGAATAAGCACTACTCTGAGCCAAGCTGAGGGAGTCCGAGCCACTTGGSA 1081
QY 1903 GCGCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGSCACT 1962
Db 1082 GCGCATTTGGGATCTTCTCATTTGATTCACTCCAGAGCACATCCAGGCTATTAAGCTGT 1141
QY 1963 TCGAGGAGAAATTTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCGCA 2022
Db 1142 CAAAGAAGAAATTTCTCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCA 1201
QY 2023 GCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAA 2082
Db 1202 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAAATTCGCCAGGCCCTGGAACAGATCAA 1261
QY 2083 CACCAGATGGAAGCTTCTGAGGTGGCGTCGAGGACCGAGTCCAGGAGTCCAGTGCATGAAGC 2142
Db 1262 CGTCCGATGGAACAACATAAGCGCTCAGTTGATGAGAGGCTTAAGCAGCTCCAGGATGC 1321
QY 2143 CCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCTTG 2202
Db 1322 CCACCGGACTTTGGGCTGGTCAAGCACTTCTCTCCTCTCTGTCCAGGTTCCCTG 1381
QY 2203 GGAGAGAGCCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAAACAC 2262
Db 1382 GGAAGAGCAATTTCAACCAATAAAGTTCCCTACTACATCAACACCAGGCTCAGACCAC 1441
QY 2263 TTGCTGGGACCATCCCAAAATGACAGAGTCTACAGTCTTTAGTGAATTAATGT 2322
Db 1442 ATGCTGGGACCATCCCAAGATGACAGATTATACCAACCCCTAGTGTCTGAACAACAT 1501
QY 2323 CAGATTCTCAGCTTATAGGACTGCCATGAAGTCCGAAAGACTCGAGAGGCTTTTGTCTT 2382
Db 1502 TAAGTTCTCAGCTTATGCACTGCCATGAAACTCCGAGAGTCCAGAAAGCCCTGCGCTT 1561
QY 2383 GGATCTCTTGA 2393

Db 1562 GGACCTGGTAA 1572

RESULT 8

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match 3.3%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 1.6e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 865 CCTAAACGCCAAGTACAAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGT 924
Db 16 CCTGAAAACCTGCTTGAAAGAACATAAAAGTTTGCAAAAGTGACCTCGAAGCTGAGCAGGT 75
QY 925 CAGGTCAATTCTCTCACTCAACATGCTGGTGGTACTTGATGAATCTAGTGGAGATCACGC 984
Db 76 GAAGTGAATTCCTTAACTCATATGTTGGTGAATTTGGATGAAACACAGTGGGAGAGCGC 135
QY 985 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGAGATCGATGGGCAAAACATCTGTAG 1044
Db 136 CACAGTGTGTTTGGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 1045 ATGGA 1049
Db 196 CTGGA 200

RESULT 9

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4
Query Match 3.3%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 2.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 853 TGATCTTGAAGACCTAAACGCCCAAGTACAAACAACATAAGGTGCTTCAAGAAGATCTAGA 912
Db 4 TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACATAAAAGTTTGCAAAATGACCTTGA 63
QY 913 ACAAGAACAAGTCAGGTCATTTCTCTCACTCAACATGCTGGTGGTAGTTGATGAATCTAG 972
Db 64 AGCTGAACAGGTGAAGTAAATTCCTTAACCTCACTGCTGGTGGTATTGTGGATGAAAACAG 123
QY 973 TGGAGATCACGCAACTGCTGCTTTTGAAGAACAACAACTTAAGGTATTGGGAGATCGATGGC 1032
Db 124 TGGGAGAGTGCCACAGCTCTTCTGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGAC 183
QY 1033 AAACATCTGTAGATGGA 1049
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 10

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match 3.3%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 2.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 853 TGATCTTGAAGACCTAAACGCCCAAGTACAAACAACATAAGGTGCTTCAAGAAGATCTAGA 912
Db 4 TGATGTGAAATCTCTACAAAAGCTGTAGAAGAACAATAAAAGTTTGCAAAAGTATCTTGA 63
QY 913 ACAAGAACAAGTCAGGTCATTTCTCTCACTCAACATGCTGGTGGTAGTTGATGAATCTAG 972
Db 64 GGCTGAACAGGTGAAGTAAATTCACCTAACTCACTGCTGGTGCATTTGTTGATGAAAACAG 123
QY 973 TGGAGATCACGCAACTGCTGCTTTTGAAGAACAACAACTTAAGGTATTGGGAGATCGATGGC 1032
Db 124 TGGTGAGAGCGCTACAGCTATCTAGAAGACCAGTTACAGAAACTTGGTGAGCGCTGGAC 183
QY 1033 AAACATCTGTAGATGGA 1049
Db 184 AGCAGTATGCCGCTTGA 200


```
Db 445 AATYAACRWAWAGAGAAWKRWKWAGSMRACAPAGTTMAWGACWAMARRGWRRCARGR--T 388
QY 679 GATGAATCTCTAAATTCAGATGGGAATGCTCAGGTTAGCTAGATGGAAAAACAAG 738
Db 387 GSWGSKGGYRMWGRMAAAKRRMAAAGSYCGMTSYTSGSKMTGRKGMTKRKMTY 328
QY 739 CAATTTACATAGAGTTTAAATGATCTCCAGATCAGAAACTGAAGAGTTGAATGACTG 798
Db 327 SGMWTSYKCTKTGKTKGWSKKTRWICTSWRKYMMWMSGCWAFSMKSWARSWYSMMAC 268
QY 799 GCTAACAACAAACAGAGAAAGAAACAAGGAATGGAGGAGCCTCTTGGACCTGATCT 858
Db 267 WCMWSASAYRARRSMYGARRSMRAGAGWWRARRGKKRARGKSSMMRSMRMSAGKA 208
QY 859 TGAAGACCTAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAGA 918
Db 207 RMCRRMWSRRMSYSWGSCKMSCRGTCAKWRYARYAKRYASSMGKYMGCRCWCYAKC 148
QY 919 ACAAGTCAGGG 929
Db 147 ARMYGYRERS 137

RESULT 14
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match 1.8%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0078;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 2201 TGGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 2260
Db 1733 TGGGAGGAAGCTTACACAGCAGAGTGAATCAAGTACTTCAATCAACACCGTGACACAGACC 1792
QY 2261 ACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAAT 2320
Db 1793 ACGTCCTGGATCCACCCCGTGATGAGCGCCCTGAACCTGTCTGTGTGACAGAGAGTGAA 1852
QY 2321 GTCAGATTCTCAGCTTATAGGACTGCCATGAAGTCCGAAG 2361
Db 1853 GAGGACTGTCCCAGAGAGCTAACAGACCCGAAAGCTGATG 1893

RESULT 15
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
```

```
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
```

NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match 1.8%; Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred.No. 0.72;
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 642 AATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGAT 701
Dy 656477 AATTAGAAGAAAGAGAAGAAATTTGAGGATATCAAGACTCAGATACAAAAT 656418
Qy 702 GGAATGCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGG 761
Dy 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCCTAGGCAAGCAATTTATGTGG 656358
Qy 762 ATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 813
Dy 656357 CAATTAAGCAGAAAAAACTAAAGCTTCTTAAAGAGACGCGCTGGGAAATAGA 656306

Search completed: April 5, 2004, 08:24:54
Job time : 121.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 04:53:30 ; Search time 582.2 Seconds
(without alignments)
15454.047 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 acatgcattcaacatcgcca.....tggatctcttgagcctgtca 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2401	100.0	3858	10	US-09-845-416-9 Sequence 9, Appli
2	2401	100.0	4825	10	US-09-845-416-29 Sequence 29, Appli
3	2401	100.0	4848	10	US-09-845-416-35 Sequence 35, Appli
4	2401	100.0	5060	10	US-09-845-416-36 Sequence 36, Appli
5	2250	93.7	3999	10	US-09-845-416-6 Sequence 6, Appli
6	2250	93.7	4966	10	US-09-845-416-28 Sequence 28, Appli
7	2250	93.7	4990	10	US-09-845-416-34 Sequence 34, Appli
8	2067	86.1	4182	10	US-09-845-416-2 Sequence 2, Appli
9	2067	86.1	5149	10	US-09-845-416-27 Sequence 27, Appli
10	1839.4	76.6	5462	15	US-10-149-736-41 Sequence 41, Appli
11	1737	72.3	3531	10	US-09-845-416-10 Sequence 10, Appli
12	1737	72.3	4498	10	US-09-845-416-30 Sequence 30, Appli
13	1725	71.8	5339	15	US-10-149-736-40 Sequence 40, Appli
14	1695	70.6	3510	10	US-09-845-416-12 Sequence 12, Appli
15	1695	70.6	4476	10	US-09-845-416-31 Sequence 31, Appli

16	1419	59.1	5417	15	US-10-149-736-39	Sequence 39, Appli
17	1335.6	55.6	8689	15	US-10-149-736-42	Sequence 42, Appli
18	1335.6	55.6	11058	10	US-09-845-416-1	Sequence 1, Appli
19	1335.6	55.6	11443	15	US-10-149-736-44	Sequence 44, Appli
20	1335.6	55.6	12057	15	US-10-149-736-47	Sequence 47, Appli
21	1335.6	55.6	13957	9	US-09-782-378A-22	Sequence 22, Appli
22	1335.6	55.6	13957	9	US-09-880-107-2284	Sequence 2284, Ap
23	1335.6	55.6	13957	15	US-10-149-736-1	Sequence 1, Appli
24	1335.6	55.6	14069	12	US-10-342-887-434	Sequence 434, App
25	1335.6	55.6	14082	12	US-10-342-887-981	Sequence 981, App
26	1335.6	55.6	14082	15	US-10-341-434-108	Sequence 108, App
27	1323	55.1	2169	10	US-09-845-416-4	Sequence 4, Appli
28	1263	52.6	4414	10	US-09-845-416-32	Sequence 32, Appli
29	1251	52.1	3446	10	US-09-845-416-14	Sequence 14, Appli
30	1143.6	47.6	13815	15	US-10-149-736-2	Sequence 2, Appli
31	1092.2	45.5	1991	10	US-09-845-416-3	Sequence 3, Appli
32	1077	44.9	1667	10	US-09-845-416-7	Sequence 7, Appli
33	976	40.6	1821	10	US-09-845-416-13	Sequence 13, Appli
34	750	31.2	1340	10	US-09-845-416-11	Sequence 11, Appli
35	589	24.5	1434	10	US-09-845-416-15	Sequence 15, Appli
36	495.6	20.6	11096	15	US-10-149-736-4	Sequence 4, Appli
37	487	20.3	10302	9	US-09-782-378A-23	Sequence 23, Appli
38	487	20.3	10302	15	US-10-149-736-3	Sequence 3, Appli
39	475	19.8	16531	14	US-10-101-510-667	Sequence 667, App
40	411.8	17.2	5106	12	US-10-220-120-157	Sequence 157, App
41	387	16.1	387	15	US-10-149-736-32	Sequence 32, Appli
42	348	14.5	348	15	US-10-149-736-31	Sequence 31, Appli
43	331	13.8	333	15	US-10-149-736-9	Sequence 9, Appli
44	327	13.6	327	15	US-10-149-736-8	Sequence 8, Appli
45	324	13.5	324	15	US-10-149-736-33	Sequence 33, Appli

ALIGNMENTS

RESULT 1
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

Query Match	100.0%;	Score 2401;	DB 10;	Length 3858;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2401;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	60	
Ds	600	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	659	
Qy	61	TGTTGATACCCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA	120	
Ds	660	TGTTGATACCCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA	719	
Qy	121	AGTTTTCCTCAACAAGTGAGCATTTGAAGSCATCCAGGAAGTGGAAATGTTGCCAAGGCC	180	
Ds	720	AGTTTTCCTCAACAAGTGAGCATTTGAAGSCATCCAGGAAGTGGAAATGTTGCCAAGGCC	779	
Qy	181	ACCTAAAGTGACTAAAGAAGACATTTTCAGTTACATCATCAATGCACTATTCTCAACA	240	

Db 780 ACCTAAAGTGACTAAAGAGAACATTTTCAGTTTACATCATCAAATGCACTATTCTCAACA 839
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 300
Db 840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 899
QY 301 GAGCTATGCCATACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAAT 360
Db 900 GAGCTATGCCATACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAAT 959
QY 361 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 420
Db 960 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 1019
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCCTTC 480
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCCTTC 1079
QY 481 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1080 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 541 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 600
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 601 TAATATTCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1200 TAATATTCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACACAGGAAATGGAGGAAGA 840
Db 1380 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACACAGGAAATGGAGGAAGA 1439
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAAGGTGCTTCA 900
Db 1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAAGGTGCTTCA 1499
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 960
Db 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1559
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
Db 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1679
QY 1081 TCATAGATTACTGCAACAGTTCCCTGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1140
Db 1680 TCATAGATTACTGCAACAGTTCCCTGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGA 1739
QY 1141 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
Db 1740 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1799
QY 1201 CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
Db 1800 CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1859
QY 1261 TCACACAGATGTTTATTCACAAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320

Db 1860 TCACACAGATGTTTATTCACAAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGA 1919
QY 1321 AGTTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
Db 1920 AGTTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1979
QY 1381 TGAACHTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
Db 1980 TGAACHTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2039
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 2040 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2099
QY 1501 ATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAACGACGATGT 1560
Db 2100 ATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAACGACGATGT 2159
QY 1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGA 1620
Db 2160 ACATAGGGCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGA 2219
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
Db 2220 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 2279
QY 1681 GCCACAGAGCTGCCTCCTGAGGAGAGAGCCACAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
Db 2280 GCCACAGAGCTGCCTCCTGAGGAGAGAGCCACAGAAATGTCACTCGGCTTCTACGAAAGCA 2339
QY 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACTCCGCTGACTGGCAGAG 1800
Db 2340 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACTCCGCTGACTGGCAGAG 2399
QY 1801 AAAAATAGATGAGACCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
Db 2400 AAAAATAGATGAGACCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2459
QY 1861 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGGATCTCCT 1920
Db 2460 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGGATCTCCT 2519
QY 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCC 1980
Db 2520 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCC 2579
QY 1981 TCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAAT 2040
Db 2580 TCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAAT 2639
QY 2041 TCAGTCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 2100
Db 2640 TCAGTCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 2699
QY 2101 GCAGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACTTTTGGTCC 2160
Db 2700 GCAGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACTTTTGGTCC 2759
QY 2161 AGCATCTCAGCACTTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
Db 2760 AGCATCTCAGCACTTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2819
QY 2221 AAACAAAGTGCCCTACTATATCAACCCAGGACTCAAACAACCTTGTGGGACCATCCCAA 2280
Db 2820 AAACAAAGTGCCCTACTATATCAACCCAGGACTCAAACAACCTTGTGGGACCATCCCAA 2879
QY 2281 AATGACAGAGCTCTACAGCTTTTAGCTGACCTGAAATAATGTGATGATTTCTCAGCTTATAG 2340
Db 2880 AATGACAGAGCTCTACAGCTTTTAGCTGACCTGAAATAATGTGATGATTTCTCAGCTTATAG 2939
QY 2341 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCTGTC 2400
Db 2940 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCTGTC 2999

2401 A 2401
3000 A 3000

RESULT 2

IS-09-845-416-29

Sequence 29, Application US/09845416

Publication No. US20030171312A1

GENERAL INFORMATION:

APPLICANT: XIAO, XIAO

TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142

CURRENT APPLICATION NUMBER: US/09/845,416

CURRENT FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 60/200,777

PRIOR FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 29

LENGTH: 4825

TYPE: DNA

ORGANISM: Homo sapiens

IS-09-845-416-29

Query Match 100.0%; Score 2401; DB 10; Length 4825;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACATGATTCACACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60

1357 ACATGATTCACACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1416

61 TGTGATPACCACCTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120

1417 TGTGATPACCACCTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476

121 AGTTTGTGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAAGTGGAAATGTTGCCAAGGCC 180

1477 AGTTTGTGCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAAGTGGAAATGTTGCCAAGGCC 1536

181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA 240

1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA 1596

241 GATCACGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 300

1597 GATCACGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA 1656

301 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGGAGCCCAT 360

1657 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGGAGCCCAT 1716

361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 420

1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGGAGAG 1776

421 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 480

1777 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 1836

481 TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540

1837 TGCTGAGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896

541 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGTTGG 600

1897 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGTTGG 1956

601 TAATATTCTACAAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA 660

1957 TAATATTCTACAAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTAATCAGAAGATGAAGA 2016

661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 720

2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 2076

721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 780

2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAAT 2136

781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 840

2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 2196

841 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAAGTACAACAACATAAAGTGCCTTCA 900

2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAAGTACAACAACATAAAGTGCCTTCA 2256

901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT 960

2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT 2316

961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACCTTAAGGTATTGGG 1020

2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACCTTAAGGTATTGGG 2376

1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAAGACAC 1080

2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAAGACAC 2436

1081 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGAAAGTTTCTTGCTGGCTTACAGA 1140

2437 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGAAAGTTTCTTGCTGGCTTACAGA 2496

1141 AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200

2497 AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 2556

1201 CTCCAAGGGAGTAAAGAGCTGTGATAAACAATGGCAAGACCTCCAAGGTGAAATGAAGC 1260

2557 CTCCAAGGGAGTAAAGAGCTGTGATAAACAATGGCAAGACCTCCAAGGTGAAATGAAGC 2616

1261 TCACACAGATGTTTATCACAACCTTGGATGAAACACAGCCCAAAAAATCCTGAGATCCCTGGA 1320

2617 TCACACAGATGTTTATCACAACCTTGGATGAAACACAGCCCAAAAAATCCTGAGATCCCTGGA 2676

1321 AGGTTCCGATGTCAGTCTCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAG 1380

2677 AGGTTCCGATGTCAGTCTCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAG 2736

1381 TGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTG 1440

2737 TGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTG 2796

1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500

2797 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2856

1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGGAGTTCAGAAAGCAGAACCGATGT 1560

2857 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGGAGTTCAGAAAGCAGAACCGATGT 2916

1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 1620

2917 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 2976

1621 GACTGTACGAATATTTCTGA CAGAGCGCTTTTGGAGGACTAGAGAAACTCTACCAGGA 1680

2977 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 3036

1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCA 1740

3037 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCA 3096

QY 1741 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1800
DB 3097 GGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 3156
QY 1801 AAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
DB 3157 AAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 3216
QY 1861 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1920
DB 3217 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3276
QY 1921 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
DB 3277 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3336
QY 1981 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
DB 3337 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 3396
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCT 2100
DB 3397 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCT 3456
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCACTGATGAAGCCACAGGGACTTTTGGTCC 2160
DB 3457 GCAGGTGGCCGTCGAGGACCGAGTCAGGCACTGATGAAGCCACAGGGACTTTTGGTCC 3516
QY 2161 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
DB 3517 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3576
QY 2221 AAACAAAAGTGCCCTACTATATCAACCACGAGACTCAACAACTTGCTGGGACCATCCCAA 2280
DB 3577 AAACAAAAGTGCCCTACTATATCAACCACGAGACTCAACAACTTGCTGGGACCATCCCAA 3636
QY 2281 AATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAG 2340
DB 3637 AATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAG 3696
QY 2341 GACTGCCATGAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 2400
DB 3697 GACTGCCATGAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 3756
QY 2401 A 2401
DB 3757 A 3757

RESULT 3
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 100.0%; Score 2401; DB 10; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
DB 1380 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1439
QY 61 TGTGATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
DB 1440 TGTGATACCACTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1499
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
DB 1500 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1559
QY 181 ACCTAAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA 240
DB 1560 ACCTAAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA 1619
QY 241 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
DB 1620 GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1679
QY 301 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCATT 360
DB 1680 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCATT 1739
QY 361 TCCCTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 420
DB 1740 TCCCTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 1799
QY 421 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAGTATTATTCGTGGCTTCTTTTC 480
DB 1800 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAGTATTATTCGTGGCTTCTTTTC 1859
QY 481 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
DB 1860 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1919
QY 541 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 600
DB 1920 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 1979
QY 601 TAATATTCTACAATTGGGAAAGTAAGCTGATTGGAAACAGGAAATAATTATCAGAAGATGAAGA 660
DB 1980 TAATATTCTACAATTGGGAAAGTAAGCTGATTGGAAACAGGAAATAATTATCAGAAGATGAAGA 2039
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 720
DB 2040 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2099
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT 780
DB 2100 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT 2159
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAATAATGGAGGAAGA 840
DB 2160 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAATAATGGAGGAAGA 2219
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAATAAGGTGCTTCA 900
DB 2220 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAATAAGGTGCTTCA 2279
QY 901 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCAGATGGTGGTAGT 960
DB 2280 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCAGATGGTGGTAGT 2339
QY 961 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1020
DB 2340 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 2399
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAC 1080
DB 2400 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAC 2459

1081	TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGA	1141
2460	TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGA	2519
1141	AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGCTCTCTAGAAGA	1200
2520	AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGCTCTCTAGAAGA	2579
1201	CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC	1260
2580	CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC	2639
1261	TCACACAGATGTTTATCAACAACCTGGATGAAACAACAGCCAAAAAATCCTGAGATCCCTGGA	1320
2640	TCACACAGATGTTTATCAACAACCTGGATGAAACAACAGCCAAAAAATCCTGAGATCCCTGGA	2699
1321	AGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG	1380
2700	AGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG	2759
1381	TGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCCAGTG	1440
2760	TGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCCAGTG	2819
1441	GAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA	1500
2820	GAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA	2879
1501	ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGACGACGATGT	1560
2880	ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGACGACGATGT	2939
1561	ACATAGGCGCTTCAAGAGGGAATTGAAACCTAAAGAACCTGTAATCATGAGTACTCTTGA	1620
2940	ACATAGGCGCTTCAAGAGGGAATTGAAACCTAAAGAACCTGTAATCATGAGTACTCTTGA	2999
1621	GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCCAGGA	1680
3000	GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCCAGGA	3059
1681	GCCCAGAGAGTGCCTCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1740
3060	GCCCAGAGAGTGCCTCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3119
1741	GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	1800
3120	GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	3179
1801	AAAAATAGATGACACCTTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA	1860
3180	AAAAATAGATGACACCTTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA	3239
1861	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCCGATCTCCT	1920
3240	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCCGATCTCCT	3299
1921	CATTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGGCC	1980
3300	CATTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGGCC	3359
1981	TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAT	2040
3360	TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAT	3419
2041	TCAGCTCTACCGTATAACCTCAGCACTCTGGAAGAAGCCTGAACACCAAGATGGAAGCTTCT	2100
3420	TCAGCTCTACCGTATAACCTCAGCACTCTGGAAGAAGCCTGAACACCAAGATGGAAGCTTCT	3479
2101	GCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCAACAGGACCTTTGGTCC	2160
3480	GCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCAACAGGACCTTTGGTCC	3539

QY	2161	AGCATCTCAGCACTTTCTTTCCACGCTCTGCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	2220
Db	3540	AGCATCTCAGCACTTTCTTTCCACGCTCTGCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	3599
QY	2221	AAACAAAGTGCCTACTATATCAACCACGAGACTCAAACTTCTGGGACCATCCAA	2280
Db	3600	AAACAAAGTGCCTACTATATCAACCACGAGACTCAAACTTCTGGGACCATCCAA	3659
QY	2281	AATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG	2340
Db	3660	AATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG	3719
QY	2341	GACTGCCATGAACTCCGAAGACTGCAGAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	2400
Db	3720	GACTGCCATGAACTCCGAAGACTGCAGAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	3779
QY	2401	A 2401	
Db	3780	A 3780	

RESULT 4

US-09-845-416-36

; Sequence 36, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 36

; LENGTH: 5060

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-36

Query Match 100.0%; Score 2401; DB 10; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 480
Db 2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 2071
QY 481 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 2072 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131
QY 541 CCAGTTTTCATACATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 600
Db 2132 CCAGTTTTCATACATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 2191
QY 601 TAATATTTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAATTTATCAGAAGATGAAGA 660
Db 2192 TAATATTTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAATTTATCAGAAGATGAAGA 2251
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 720
Db 2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 2311
QY 721 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAAAAT 780
Db 2312 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAAAAT 2371
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAAATGAGGAAGA 840
Db 2372 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAAATGAGGAAGA 2431
QY 841 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACATTAAGTGTCTTCA 900
Db 2432 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACATTAAGTGTCTTCA 2491
QY 901 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 960
Db 2492 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2551
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGTATTGGG 1020
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGTATTGGG 2611
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAAGACAC 1080
Db 2612 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAAGACAC 2671
QY 1081 TCATAGATTACTGCAACAGTTCCCTCCCTGGACCTTGGAAAAGTTCTTTCCTGGCTTACAGA 1140
Db 2672 TCATAGATTACTGCAACAGTTCCCTCCCTGGACCTTGGAAAAGTTCTTTCCTGGCTTACAGA 2731
QY 1141 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCGTAAGGAAAGGCTCCTAGAAGA 1200
Db 2732 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCGTAAGGAAAGGCTCCTAGAAGA 2791
QY 1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1260
Db 2792 CTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 2851
QY 1261 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATAATCCTGAGATCCCTTGA 1320
Db 2852 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATAATCCTGAGATCCCTTGA 2911
QY 1321 AGGTTCCGATGATGCAGTCCCTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
Db 2912 AGGTTCCGATGATGCAGTCCCTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2971
QY 1381 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 1440
Db 2972 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG 3031
QY 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 3032 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091

QY 1501 ATTAAGCCGGCAGSCACCTATTGGAGGGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 1560
Db 3092 ATTAAGCCGGCAGSCACCTATTGGAGGGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGT 3151
QY 1561 ACATAGGGCCCTTCAAGAGGGAATTGAAAACCTTAAAGAACCTGTAAATCATGACTACTCTTGA 1620
Db 3152 ACATAGGGCCCTTCAAGAGGGAATTGAAAACCTTAAAGAACCTGTAAATCATGACTACTCTTGA 3211
QY 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAAATCTTACCAGGA 1680
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAAATCTTACCAGGA 3271
QY 1681 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCCAGAAATGTCTACTCGGCTTCTACGAAAGCA 1740
Db 3272 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCCAGAAATGTCTACTCGGCTTCTACGAAAGCA 3331
QY 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAATTTGAAACCTGCACTCCGCTGACTGGCAGAG 1800
Db 3332 GGCTGAGGAGGTCAATACTGAGTGGGAAAATTTGAAACCTGCACTCCGCTGACTGGCAGAG 3391
QY 1801 AAAAATAGATGAGACCCCTTTGAAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
Db 3392 AAAAATAGATGAGACCCCTTTGAAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 3451
QY 1861 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1920
Db 3452 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3511
QY 1921 CATTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGCCCC 1980
Db 3512 CATTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAAATTTGCCCC 3571
QY 1981 TCTGAAAGAGAACCTGAGGCCACGTCAATGACCTTGGAAGACCTGAACACAGATGGAAGCTTCT 2040
Db 3572 TCTGAAAGAGAACCTGAGGCCACGTCAATGACCTTGGAAGACCTGAACACAGATGGAAGCTTCT 3631
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCT 2100
Db 3632 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCT 3691
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCC 2160
Db 3692 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCC 3751
QY 2161 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
Db 3752 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3811
QY 2221 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGAGCCATCCCAA 2280
Db 3812 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGAGCCATCCCAA 3871
QY 2281 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCTTATAG 2340
Db 3872 AATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCTTATAG 3931
QY 2341 GACTGCCATGAAACTCCGAAGACTGAGAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTC 2400
Db 3932 GACTGCCATGAAACTCCGAAGACTGAGAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTC 3991
QY 2401 A 2401
Db 3992 A 3992

RESULT 5
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF

FILE REFERENCE: DB1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
JS-09-845-416-6

Query Match 93.7%; Score 2250; DB 10; Length 3999;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

1 ACATGATTCACATCGCCAGATATCAATTAGGATAGAGAAACTACTCGATCCTGAAGA 60
|||||
600 ACATGATTCACATCGCCAGATATCAATTAGGATAGAGAAACTACTCGATCCTGAAGA 659
|||||
61 TGTGATPACACCTATCCAGATPAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
|||||
660 TGTGATPACACCTATCCAGATPAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA 719
|||||
121 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
|||||
720 AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
|||||
181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTTCTCAACA 240
|||||
780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTTCTCAACA 839
|||||
241 GATCACGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
|||||
840 GATCACGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 899
|||||
301 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACAGGAGCCCATTT 360
|||||
900 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACAGGAGCCCATTT 959
|||||
361 TCCTTCACAGCATTTTGGAGGCTCCTGAAGACAAAGTTCATTTGGCAGTTTATTGATGGAGAG 420
|||||
960 TCCTTCACAGCATTTTGGAGGCTCCTGAAGACAAAGTTCATTTGGCAGTTTATTGATGGAGAG 1019
|||||
421 TGAAGTAAACCTGGACCGGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 480
|||||
1020 TGAAGTAAACCTGGACCGGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 1079
|||||
481 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 540
|||||
1080 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
|||||
541 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 600
|||||
1140 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1199
|||||
601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 660
|||||
1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1259
|||||
661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 720
|||||
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 1319
|||||
721 TAGCATGGAAAAACAAAGCAATTTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 780
|||||
1320 TAGCATGGAAAAACAAAGCAATTTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 1379
|||||
781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAAAGGAAAAATGGAGGAAGA 840
|||||
1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAAAGGAAAAATGGAGGAAGA 1439
|||||

841 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 900
|||||
1440 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1499
|||||
901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 960
|||||
1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1559
|||||
961 TGATGAATCTACTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
|||||
1560 TGATGAATCTACTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1619
|||||
1021 AGATCGATGGSCAAACATCTGTAGATGGACAGAACCGCTGGGTCTCTTTTACAGAC-- 1078
|||||
1620 AGATCGATGGSCAAACATCTGTAGATGGACAGAACCGCTGGGTCTCTTTTACAGACCA 1679
|||||
1079 ----- 1078
|||||
1680 GCCTGACCTAGCTCCTGGACTGACCACACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 1739
|||||
1079 ----- 1078
|||||
1740 GGTGACACAACCTGTGGTTACTAAGAAACTGCCATCTCCAAACTAGAAATGCCATCTTTC 1799
|||||
1079 ----- ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAA 1119
|||||
1800 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAA 1859
|||||
1120 GTTTCTTGCCTGGCTTACAGAAAGTGAACAACTGCCAATGTCTCTACAGGATGCTACCCG 1179
|||||
1860 GTTTCTTGCCTGGCTTACAGAAAGTGAACAACTGCCAATGTCTCTACAGGATGCTACCCG 1919
|||||
1180 TAAGGAAAGGCTCTAGAAGACTTCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 1239
|||||
1920 TAAGGAAAGGCTCTAGAAGACTTCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 1979
|||||
1240 CCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAAGCCCA 1299
|||||
1980 CCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAAGCCCA 2039
|||||
1300 AAAAACTCTGAGATCCCTGGAAGTTCGATGATGCGAGTCTGTTACAAAAGACGTTTGA 1359
|||||
2040 AAAAACTCTGAGATCCCTGGAAGTTCGATGATGCGAGTCTGTTACAAAAGACGTTTGA 2099
|||||
1360 TAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 1419
|||||
2100 TAACATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTT 2159
|||||
1420 GGAAGCCAGTCTGACCAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTCTCGGTGTG 1479
|||||
2160 GGAAGCCAGTCTGACCAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTCTCGGTGTG 2219
|||||
1480 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 1539
|||||
2220 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGC 2279
|||||
1540 AGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTTCAGAGGGGAATTGAAAACTAAAGAAC 1599
|||||
2280 AGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTTCAGAGGGGAATTGAAAACTAAAGAAC 2339
|||||
1600 TGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGG 1659
|||||
2340 TGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGG 2399
|||||
1660 ACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGT 1719
|||||
2400 ACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGT 2459
|||||
1720 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 1779
|||||
2460 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 2519
|||||
1780 GCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCA 1839
|||||

Db 2520 GCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCA 2579
QY 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTG 1899
Db 2580 AGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTG 2639
QY 1900 GCAGCCCGTGGCGATCTCCTCATTTGACCTCTCCAAGATCACTCGAGAAAGTCAAGGC 1959
Db 2640 GCAGCCCGTGGCGATCTCCTCATTTGACCTCTCCAAGATCACTCGAGAAAGTCAAGGC 2699
QY 1960 ACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 2019
Db 2700 ACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 2759
QY 2020 CCAGCTTACCACTTTGGGCATTTCAGCTCTCACCCTATTAACCTCAGCACTCTGGAAGACCT 2079
Db 2760 CCAGCTTACCACTTTGGGCATTTCAGCTCTCACCCTATTAACCTCAGCACTCTGGAAGACCT 2819
QY 2080 GAAACACCAGATGGAAGCTTCTGACGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGA 2139
Db 2820 GAAACACCAGATGGAAGCTTCTGACGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGA 2879
QY 2140 AGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCC 2199
Db 2880 AGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCC 2939
QY 2200 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAAC 2259
Db 2940 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAAC 2999
QY 2260 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 2319
Db 3000 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3059
QY 2320 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 2379
Db 3060 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 3119
QY 2380 CTTGGATCTCTTGAGCCTGTCA 2401
Db 3120 CTTGGATCTCTTGAGCCTGTCA 3141

RESULT 6

US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 93.7%; Score 2250; DB 10; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1416

QY 61 TGTGTATACCACCTATCCAGATAAAGAGTCCATCTTAATGTACATCACATCACTCTTTCCA 120
Db 1417 TGTGTATACCACCTATCCAGATAAAGAGTCCATCTTAATGTACATCACATCACTCTTTCCA 1476
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 1477 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536
QY 181 ACCTAAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 240
Db 1537 ACCTAAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 1596
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
Db 1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656
QY 301 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 360
Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1716
QY 361 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 420
Db 1717 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1776
QY 421 TGAATAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATATCGTGGCTTCTTTTC 480
Db 1777 TGAATAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATATCGTGGCTTCTTTTC 1836
QY 481 TGCTGAGGACACATTTGAAAGCAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 540
Db 1837 TGCTGAGGACACATTTGAAAGCAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 1896
QY 541 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 600
Db 1897 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
Db 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGAAATGGAGGAAGA 840
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGAAATGGAGGAAGA 2196
QY 841 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAAACATAAGGTGCTTCA 900
Db 2197 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAAACATAAGGTGCTTCA 2256
QY 901 AGAAGATCTAGAACAAAGAAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 960
Db 2257 AGAAGATCTAGAACAAAGAAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 2316
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
QY 1021 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGAC-- 1078
Db 2377 AGATCGATGGGCAACACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACCA 2436
QY 1079 ----- 1078
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 2496
QY 1079 ----- 1078

2497 GGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 2556
1079 -----ACTCATAGATTACTGCAACAGATTCCCTCGACCTGGAAAA 1119
2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGATTCCCTCGACCTGGAAAA 2616
1120 GTTTCCTGCTGGCTTACAGAAAGTGAACAACTGCCAATGTCTTACAGGATGCTACCCG 1179
2617 GTTTCCTGCTGGCTTACAGAAAGTGAACAACTGCCAATGTCTTACAGGATGCTACCCG 2676
1180 TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGTATGAAACAAATGGCAAGA 1239
2677 TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGTATGAAACAAATGGCAAGA 2736
1240 CCTCCAAGGTGAAATTAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCA 1299
2737 CCTCCAAGGTGAAATTAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCA 2796
1300 AAAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTCTGTATACAAAGCGTTTGA 1359
2797 AAAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTCTGTATACAAAGCGTTTGA 2856
1360 TAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTT 1419
2857 TAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTCCCATTT 2916
1420 GGAAGCCAGTCTGACCAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTCTGGTGTG 1479
2917 GGAAGCCAGTCTGACCAGTGGAAAGCTCTGCACCTTCTCTGCAGGAACCTCTGGTGTG 2976
1480 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGCGGACTTCCAGC 1539
2977 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGCGGACTTCCAGC 3036
1540 AGTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAACCTAAAGAAC 1599
3037 AGTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAACCTAAAGAAC 3096
1600 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTGGAAG 1659
3097 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTGGAAG 3156
1660 ACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGT 1719
3157 ACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGT 3216
1720 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 1779
3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 3276
1780 GCACCTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCA 1839
3277 GCACCTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCA 3336
1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 1899
3337 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 3396
1900 GCAGCCCGTGGGCGATCTCCTAATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGC 1959
3397 GCAGCCCGTGGGCGATCTCCTAATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGC 3456
1960 ACTTCGAGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 2019
3457 ACTTCGAGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 3516
2020 CCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCT 2079
3517 CCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCT 3576
2080 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGAGCCGAGTCAAGGAGCTGCATGA 2139

3577 GAACACCAGATGGAAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGA 3636
2140 AGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCC 2199
3637 AGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCC 3696
2200 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAA 2259
3697 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAA 3756
2260 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAA 2319
3757 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAA 3816
2320 TGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGCTGCAGAAAGGCCCTTG 2379
3817 TGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGCTGCAGAAAGGCCCTTG 3876
2380 CTTGGATCTCTTGAGCCCTGTCA 2401
3877 CTTGGATCTCTTGAGCCCTGTCA 3898

RESULT 7

US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 93.7%; Score 2250; DB 10; Length 4990;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 ACATGCAATTCACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
D 1381 ACATGCAATTCACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 1440
QY 61 TGTGATACCACTATCCAGATAAGAGTCCCATCTTAATGTACATCACATCACATCTTTCCA 120
D 1441 TGTGATACCACTATCCAGATAAGAGTCCCATCTTAATGTACATCACATCACATCTTTCCA 1500
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAAAATGTTGCCAAGGC 180
D 1501 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAAAATGTTGCCAAGGC 1560
QY 181 ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTTACATCATCAAAATGCACATATTCTCAACA 240
D 1561 ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTTACATCATCAAAATGCACATATTCTCAACA 1620
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 300
D 1621 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA 1680
QY 301 GAGCTATGCCCTACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 360
D 1681 GAGCTATGCCCTACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 1740
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 420

Db 1741 TCCTTACAGCATTTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTTGATGGAGAG 1800
Qy 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 480
Db 1801 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 1860
Qy 481 TGCTGAGGACACATTGCAAGCACAAGGAGACATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1861 TGCTGAGGACACATTGCAAGCACAAGGAGACATTTCTAATGATGTGGAAGTGGTGAAGA 1920
Qy 541 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTGG 600
Db 1921 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTGG 1980
Qy 601 TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1981 TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2040
Qy 661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
Db 2041 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2100
Qy 721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 2101 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2160
Qy 781 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 840
Db 2161 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 2220
Qy 841 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAAACATAAGGTGCTTCA 900
Db 2221 GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAAACATAAGGTGCTTCA 2280
Qy 901 AGAAGATCTAGAACAAAGACAAAGTCAGGGTCAATTTCTCACTCAATGGTGGTAGT 960
Db 2281 AGAAGATCTAGAACAAAGACAAAGTCAGGGTCAATTTCTCACTCAATGGTGGTAGT 2340
Qy 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 2341 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2400
Qy 1021 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGAC-- 1078
Db 2401 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACCA 2460
Qy 1079 ----- 1078
Db 2461 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 2520
Qy 1079 ----- 1078
Db 2521 GGTGACACAACCTGTGTGTTACTAAGGAAACTGCCATCTCCAACTAGAAATGCCATCTTC 2580
Qy 1079 -----ACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAA 1119
Db 2581 CTTGATGTTGAGGTACCTACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAA 2640
Qy 1120 GTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCAATGTCTTACAGATGCTACCCG 1179
Db 2641 GTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCAATGTCTTACAGATGCTACCCG 2700
Qy 1180 TAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGA 1239
Db 2701 TAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGA 2760
Qy 1240 CCTCCAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCA 1299
Db 2761 CCTCCAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCA 2820
Qy 1300 AAAAATCCTGAGATCCCTGGAGGTTCCGATGATGAGTCCCTGTTTACAAGACGTTTGG 1359

Db 2821 AAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTTGGA 2880
Qy 1360 TAAATGAACCTTCAAGTGGAGTGAACTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTT 1419
Db 2881 TAAATGAACCTTCAAGTGGAGTGAACTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTT 2940
Qy 1420 GGAAGCCAGTTCTGACCAAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 1479
Db 2941 GGAAGCCAGTTCTGACCAAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 3000
Qy 1480 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCCAGC 1539
Db 3001 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCCAGC 3060
Qy 1540 AGTTCAAGACAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTAAAGAACC 1599
Db 3061 AGTTCAAGACAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTAAAGAACC 3120
Qy 1600 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGG 1659
Db 3121 TGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGG 3180
Qy 1660 ACTAGAGAAACTCTACAGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATGT 1719
Db 3181 ACTAGAGAAACTCTACAGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATGT 3240
Qy 1720 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCT 1779
Db 3241 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCT 3300
Qy 1780 GCCTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCA 1839
Db 3301 GCCTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCA 3360
Qy 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGGATCCTG 1899
Db 3361 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGGATCCTG 3420
Qy 1900 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGC 1959
Db 3421 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGC 3480
Qy 1960 ACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGCCACCTCAATGACCTTTGCTCG 2019
Db 3481 ACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGCCACCTCAATGACCTTTGCTCG 3540
Qy 2020 CCAGCTTACCCTTTGGGCAATTCAGCTCTCAACGATATAAAGCTCAGCACTCTGGAAGACCT 2079
Db 3541 CCAGCTTACCCTTTGGGCAATTCAGCTCTCAACGATATAAAGCTCAGCACTCTGGAAGACCT 3600
Qy 2080 GAACACCAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCAGGAGCTGCATGA 2139
Db 3601 GAACACCAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCAGGAGCTGCATGA 3660
Qy 2140 AGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 2199
Db 3661 AGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 3720
Qy 2200 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAC 2259
Db 3721 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAC 3780
Qy 2260 AACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAA 2319
Db 3781 AACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAA 3840
Qy 2320 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCAGAAAGGCCCTTTG 2379
Db 3841 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCAGAAAGGCCCTTTG 3900
Qy 2380 CTTGGATCTCTGAGCCTGTCA 2401
Db 3901 CTTGGATCTCTGAGCCTGTCA 3922

RESULT 8
JS-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-09-845-416-2

Query Match 86.1%; Score 2067; DB 10; Length 4182;
Best local Similarity 88.1%; Pred. No: 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

2y	1	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	60
2b	600	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	659
2y	61	TGTTGATACACCTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA	120
2b	660	TGTTGATACACCTATCCAGATAAGAAAGTCCATCTTAATGTACATCACATCACTCTTCCA	719
2y	121	AGTTTTCCTCAACAAGTGAGCAITGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC	180
2b	720	AGTTTTCCTCAACAAGTGAGCAITGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC	779
2y	181	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA	240
2b	780	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA	839
2y	241	GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA	300
2b	840	GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTTAAGCCTCGATTCAA	899
2y	301	GAGCTATGCCATACACACAGGCTGTTATGTACCACTGTGACCCCTACACGGAGCCCAT	360
2b	900	GAGCTATGCCATACACACAGGCTGTTATGTACCACTGTGACCCCTACACGGAGCCCAT	959
2y	361	TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG	420
2b	960	TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG	1019
2y	421	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC	480
2b	1020	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC	1079
2y	481	TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA	540
2b	1080	TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA	1139
2y	541	CCAGTTTCACTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	600
2b	1140	CCAGTTTCACTACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG	1199
2y	601	TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAAATTTATCAGAGATGAAGA	660
2b	1200	TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAAATTTATCAGAGATGAAGA	1259
2y	661	AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	720
2b			

Db	1260	AACTGAAGTACAAGACAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC	1319
Qy	721	TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATCAGAAACT	780
Db	1320	TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATCAGAAACT	1379
Qy	781	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA	840
Db	1380	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAACAAAGGAAAATGGAGGAAGA	1439
Qy	841	GCCTCTTTGGACCTGATCTTTGAAGACCTTAAAAACGCCAAGTACAAACATAAAGTGCTTCA	900
Db	1440	GCCTCTTTGGACCTGATCTTTGAAGACCTTAAAAACGCCAAGTACAAACATAAAGTGCTTCA	1499
Qy	901	AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT	960
Db	1500	AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT	1559
Qy	961	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG	1020
Db	1560	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG	1619
Qy	1021	AGATCGATGGGCAAAACATCTGTAGATGACAGAAAGACCGCTGGTTCCTTTTACAAGAC--	1078
Db	1620	AGATCGATGGGCAAAACATCTGTAGATGACAGAAAGACCGCTGGTTCCTTTTACAAGACAT	1679
Qy	1079	-----	1078
Db	1680	CCTTCTCAAATGGCAACGTCTTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTCAGA	1739
Qy	1079	-----	1078
Db	1740	AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT	1799
Qy	1079	-----	1078
Db	1800	ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT	1859
Qy	1079	-----	1078
Db	1860	GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC	1919
Qy	1079	-----	1078
Db	1920	CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA	1979
Qy	1079	-----	1116
Db	1980	ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA	2039
Qy	1117	AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAAACAACTGCCAATCTCTACAGGATGCTAC	1176
Db	2040	AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAAACAACTGCCAATCTCTACAGGATGCTAC	2099
Qy	1177	CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA	1236
Db	2100	CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA	2159
Qy	1237	AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTTATCACAACCTGGATGAAAACAG	1296
Db	2160	AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTTATCACAACCTGGATGAAAACAG	2219
Qy	1297	CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT	1356
Db	2220	CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT	2279
Qy	1357	GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAGTCTCTCAACATTAGGTCCCA	1416
Db	2280	GGATAACATGAACITCAAGTGGAGTGAACITTCGAAAAAGTCTCTCAACATTAGGTCCCA	2339
Qy	1417	TTTGAAGCCAGTTTCTGACCCAGTGGAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGTT	1476
Db	2340	TTTGAAGCCAGTTTCTGACCCAGTGGAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGTT	2399

Y 841 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCAAGTACAAACATAAAGTGCTTCA 900
b 2197 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACCGCAAGTACAAACATAAAGTGCTTCA 2256
Y 901 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTACTCACATGGTGGTGGT 960
b 2257 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTACTCACATGGTGGTGGT 2316
Y 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1020
b 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 2376
Y 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGAC-- 1078
b 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 2436
Y 1079 ----- 1078
b 2437 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTCAGA 2496
Y 1079 ----- 1078
b 2497 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 2556
Y 1079 ----- 1078
b 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2616
Y 1079 ----- 1078
b 2617 GGGCAAACTGTATTCACTCAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2676
Y 1079 ----- 1078
b 2677 CCAGAAGACGGAAGCATGGCTGGATAACITTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2736
Y 1079 -----ACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTTGA 1116
b 2737 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTTGA 2796
Y 1117 AAAGTTTCTTGCTGGCTTACAGAAAGTGAACAACTGCCAATGCTCTACAGGATGCTAC 1176
b 2797 AAAGTTTCTTGCTGGCTTACAGAAAGTGAACAACTGCCAATGCTCTACAGGATGCTAC 2856
Y 1177 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGGCA 1236
b 2857 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGGCA 2916
Y 1237 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTTATCAACACCTGGATGAACACAG 1296
b 2917 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTTATCAACACCTGGATGAACACAG 2976
Y 1297 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTTACAAGACGTTT 1356
b 2977 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTTACAAGACGTTT 3036
Y 1357 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1416
b 3037 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3096
Y 1417 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGCTGACACCTTTCTCTGACGGAACCTTCTGTT 1476
b 3097 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGCTGACACCTTTCTCTGACGGAACCTTCTGTT 3156
Y 1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 1536
b 3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 3216
Y 1537 ACCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGGAATTGAACATAAAGA 1596
b 3217 ACCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGGAATTGAACATAAAGA 3276

QY 1597 ACCTGTAAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTTGA 1656
Db 3277 ACCTGTAAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTTGA 3336
QY 1657 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA 1716
Db 3337 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA 3396
QY 1717 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1776
Db 3397 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 3456
QY 1777 CCTGCACTCCGCTGACTGTCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1836
Db 3457 CCTGCACTCCGCTGACTGTCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3516
QY 1837 TCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1896
Db 3517 TCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3576
QY 1897 CTGGAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1956
Db 3577 CTGGAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3636
QY 1957 GGCACCTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2016
Db 3637 GGCACCTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3696
QY 2017 TCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGA 2076
Db 3697 TCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGA 3756
QY 2077 CCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCA 2136
Db 3757 CCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCA 3816
QY 2137 TGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTGTCTCCAGGG 2196
Db 3817 TGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTGTCTCCAGGG 3876
QY 2197 TCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCA 2256
Db 3877 TCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCA 3936
QY 2257 AACAACTTGCTGGGACCATCCCAAAATGACAGAGTCCATGAAACTCCGAAGACTGAGAAAGGCCCT 2376
Db 3937 AACAACTTGCTGGGACCATCCCAAAATGACAGAGTCCATGAAACTCCGAAGACTGAGAAAGGCCCT 3996
QY 2317 TAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGAGAAAGGCCCT 2401
Db 3997 TAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGAGAAAGGCCCT 4056
QY 2377 TTGCTTGATCTCTTGAGCCCTGTCA 2401
Db 4057 TTGCTTGATCTCTTGAGCCCTGTCA 4081

RESULT 10
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

```

Query Match	76.6%;	Score 1839.4;	DB 15;	Length 5462;
Best Local Similarity	88.1%;	Pred. No. 0;		
Matches 2115;	Conservative	0;	Mismatches	76;
			Indels	210;
			Gaps	2;

2Y	1	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	60
Db	799	ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	858
QY	61	TGTTGATACCACTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA	120
Db	859	TGTTGATACCACTATCCAGATAAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA	918
QY	121	AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGCAAGTGGAAATGTTGCAAGGCC	180
Db	919	AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGCAAGTGGAAATGTTGCAAGGCC	978
QY	181	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA	240
Db	979	ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAAATGCATATTCTCAACA	1038
QY	241	GATCAGGTCAGTCTAGCACAGGGATATGAGAACTTCTTCCCCTAAGCCTCGATTCAA	300
Db	1039	GATCAGGTCAGTCTAGCACAGGGATATGAGAACTTCTTCCCCTAAGCCTCGATTCAA	1098
QY	301	GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT	360
Db	1099	GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT	1158
QY	361	TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	420
Db	1159	TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	1218
QY	421	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGCCTCTTTC	480
Db	1219	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGCCTCTTTC	1278
QY	481	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	540
Db	1279	TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338
QY	541	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGGTTGG	600
Db	1339	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGGTTGG	1398
QY	601	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA	660
Db	1399	TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA	1458
QY	661	AACCTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC	720
Db	1459	AACCTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC	1518
QY	721	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAACT	780
Db	1519	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAACT	1558
QY	781	GAAAGAGTTGAATGACTGGCTPACAAAAACAGAAGAAAGAACAGGAAAATGGAGGAAGA	840
Db	1559	-----AC	1560
QY	841	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA	900
Db	1561	CACATATTGGAGCCTCTCCTACTCAGACTGTTAATCTGGTGACACAACCTGGTTACTAA	1620

Db 2569 TCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACITTTGGGCAT 2628
QY 2041 TCAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCT 2100
Db 2629 TCAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCT 2688
QY 2101 GCAGGTGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTCC 2160
Db 2689 GCAGGTGCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTCC 2748
QY 2161 AGCATCTCAGCACTTTCTTTCCACGTCGTGCCAGGTGCCGAGGAGAGGCCATCTCGCC 2220
Db 2749 AGCATCTCAGCACTTTCTTTCCACGTCGTGCCAGGTGCCGAGGAGAGGCCATCTCGCC 2808
QY 2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAACAACTTGTGGGACCATCCCAA 2280
Db 2809 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAACAACTTGTGGGACCATCCCAA 2868
QY 2281 AATGACAGAGCTCTACCACTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
Db 2869 AATGACAGAGCTCTACCACTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2928
QY 2341 GACTGCCATGAAGTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db 2929 GACTGCCATGAAGTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2988
QY 2401 A 2401
Db 2989 A 2989

RESULT 11
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-09-845-416-10

Query Match 72.3%; Score 1737; DB 10; Length 3531;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2074; Conservative 0; Mismatches 0; Indels 327; Gaps 1;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
Db 600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659
QY 61 TGTGTATACCACTATCCAGATAAGAGTCCCATCTTAATGTACATCACATCACTCTTCCA 120
Db 660 TGTGTATACCACTATCCAGATAAGAGTCCCATCTTAATGTACATCACATCACTCTTCCA 719
QY 121 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 720 AGTTTGGCTCAACAAGTGAGCATTTGAAGCCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
QY 181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 240
Db 780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAATGCACTATTCTCAACA 839

QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 300
Db 840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAAGCCTCGATTCAA 899
QY 301 GAGCTATGCCTACACACAGGCTGCTTATGTATGTCACCACTCTGACCTACACGGAGCCCAT 360
Db 900 GAGCTATGCCTACACACAGGCTGCTTATGTATGTCACCACTCTGACCTACACGGAGCCCAT 959
QY 361 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCACTTCAATTTGATGGAGAG 420
Db 960 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCACTTCAATTTGATGGAGAG 1019
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
QY 481 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1080 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 541 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 600
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1199
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAATTATCAGAAGATGAAGA 660
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAATTATCAGAAGATGAAGA 1259
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGAAATGCCTCAGGGTAGC 720
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGAAATGCCTCAGGGTAGC 1319
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGA----- 1350
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAACAGGAAAAATGGAGGAAGA 840
Db 1351 ----- 1350
QY 841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAAGTGCTTCA 900
Db 1351 ----- 1350
QY 901 AGAAGATCTAGAACACAGAACAAGTCAAGGTCOAATTTCTCACTCACATGGTGGTAGT 960
Db 1351 ----- 1350
QY 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACATAAGGTATTGGG 1020
Db 1351 ----- 1350
QY 1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
Db 1351 -----AC 1352
QY 1081 TCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1140
Db 1353 TCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1412
QY 1141 AGGTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 1200
Db 1413 AGGTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 1472
QY 1201 CTCGAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 1260
Db 1473 CTCGAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGC 1532
QY 1261 TCACACAGATGTTTTATCAAACTGGATGAACCAAGCAGCCAAAATTCCTGAGATCCCTGGA 1320
Db 1533 TCACACAGATGTTTTATCAAACTGGATGAACCAAGCAGCCAAAATTCCTGAGATCCCTGGA 1592
QY 1321 AGGTTCCGATGATGCACTCTCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380

Db 1593 AGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1652
QY 1381 TGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAAGTG 1440
Db 1653 TGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAAGTG 1712
QY 1441 GAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 1713 GAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGA 1772
QY 1501 ATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1560
Db 1773 ATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGT 1832
QY 1561 ACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTGA 1620
Db 1833 ACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTGA 1892
QY 1621 GACTGTACGAATATTTCTGTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 1680
Db 1893 GACTGTACGAATATTTCTGTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 1952
QY 1681 GCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
Db 1953 GCCAGAGAGCTGCCCTCTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2012
QY 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTTGCACTCCGCTGACTGGCAGAG 1800
Db 2013 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTTGCACTCCGCTGACTGGCAGAG 2072
QY 1801 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAATTTCAAGAGGCCACCGATGAGCTGGA 1860
Db 2073 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAATTTCAAGAGGCCACCGATGAGCTGGA 2132
QY 1861 CCTAAGCTGCGCCAAAGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCGATCTCCT 1920
Db 2133 CCTAAGCTGCGCCAAAGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCGATCTCCT 2192
QY 1921 CATTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCC 1980
Db 2193 CATTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCC 2252
QY 1981 TCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTGCCAGCTTACACCTTTGGGCAT 2040
Db 2253 TCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTGCCAGCTTACACCTTTGGGCAT 2312
QY 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCCAGATGGAAGCTTCT 2100
Db 2313 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCCAGATGGAAGCTTCT 2372
QY 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGACTTTGGTCC 2160
Db 2373 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGACTTTGGTCC 2432
QY 2161 AGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2220
Db 2433 AGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 2492
QY 2221 AAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAA 2280
Db 2493 AAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGCTGGGACCATCCCAA 2552
QY 2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGTTCTCAGCTTATAG 2340
Db 2553 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGAGTTCTCAGCTTATAG 2612
QY 2341 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 2400
Db 2613 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTC 2672
QY 2401 A 2401

Db 2673 A 2673
RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30
Query Match 72.3%; Score 1737; DB 10; Length 4498;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2074; Conservative 0; Mismatches 0; Indels 327; Gaps 1;
QY 1 ACATGCAATCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
Db 1357 ACATGCAATCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1416
QY 61 TGTGTATACACACCTATCCAGATAAAGAGTCCATCTTAAATGTACATCACATCACTCTTCCA 120
Db 1417 TGTGTATACACACCTATCCAGATAAAGAGTCCATCTTAAATGTACATCACATCACTCTTCCA 1476
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGAAAGTGGAAATGTTGCCAAGGCC 180
Db 1477 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGAAAGTGGAAATGTTGCCAAGGCC 1536
QY 181 ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 240
Db 1537 ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACATTTCTCAACA 1596
QY 241 GATCACGTCAGTCTAGCACAGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 300
Db 1597 GATCACGTCAGTCTAGCACAGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 1656
QY 301 GAGCTATGCTTACACACAGGCTGTTATGTCAACACCTTGACCCCTACACGGAGCCCAT 360
Db 1657 GAGCTATGCTTACACACAGGCTGTTATGTCAACACCTTGACCCCTACACGGAGCCCAT 1716
QY 361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG 420
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG 1776
QY 421 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAATTTATCGTGGCTTCTTTC 480
Db 1777 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAATTTATCGTGGCTTCTTTC 1836
QY 481 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTTCTAATGATGTGGAAGTGGTGAAGA 540
Db 1837 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 541 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 600
Db 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1956
QY 601 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 660
Db 1957 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 661 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAGGTAGC 720

Db 2017 AACTGAAGTACAAAGACGAGATGATCTCCTAAATTC AAGATGGGAATGCCTCAGGTTAGC 2076
2y 721 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 780
Db 2077 TAGCATGGAAAAACAAAGCAATTTTACATAGA----- 2107
2y 781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 840
Db 2108 ----- 2107
2y 841 GCCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA 900
Db 2108 ----- 2107
2y 901 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 960
Db 2108 ----- 2107
Qy 961 TGATGAATCTAGTGGAGATCAACGCAACTGCTGTCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db 2108 ----- 2107
Qy 1021 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAAGACAC 1080
Db 2108 -----AC 2109
Qy 1081 TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCTGGCTTACAGA 1140
Db 2110 TCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCTGGCTTACAGA 2169
2y 1141 AGCTGAAACAACTGCCAATGTCTTACAGATGCTACCCGTGAAGAAAGGCTCCCTAGAAGA 1200
Db 2170 AGCTGAAACAACTGCCAATGTCTTACAGATGCTACCCGTGAAGAAAGGCTCCCTAGAAGA 2229
2y 1201 CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1260
Db 2230 CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 2289
2y 1261 TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAATAATCCTGAGATCCCTGGA 1320
Db 2290 TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAATAATCCTGAGATCCCTGGA 2349
Qy 1321 AGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTG 1440
Db 2410 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTG 2469
2y 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAATTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
Db 2470 GAAGCGTCTGCACCTTTCTCTGCAGGAATTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2529
2y 1501 ATTAAGCCGGCAGGCACCTATTGGAGGGGACTTTCCAGCAGTTTCAGAAAGCAGAACCATGT 1560
Db 2530 ATTAAGCCGGCAGGCACCTATTGGAGGGGACTTTCCAGCAGTTTCAGAAAGCAGAACCATGT 2589
2y 1561 ACATAGGGCCTTCAAGAGGGGAATTTGAAAACTAAAGAACCTGTGAATCATGAGTACTCTTGA 1620
Db 2590 ACATAGGGCCTTCAAGAGGGGAATTTGAAAACTAAAGAACCTGTGAATCATGAGTACTCTTGA 2649
Qy 1621 GACTGTACGAATATTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 1680
Db 2650 GACTGTACGAATATTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 2709
Qy 1681 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTCTACGAAGCA 1740
Db 2710 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTCTACGAAGCA 2769
Qy 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTTGCACTCCGCTGACTGGCAGAG 1800

Db 2770 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2829
Qy 1801 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
Db 2830 AAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2889
Qy 1861 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGAGCCCGTGGCGGATCTCCT 1920
Db 2890 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGAGCCCGTGGCGGATCTCCT 2949
Qy 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGGCC 1980
Db 2950 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGGCC 3009
Qy 1981 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTTGGGCAT 2040
Db 3010 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTTGGGCAT 3069
Qy 2041 TCAGCTCTCACCGTATAACCTCAGCACCTCGGAAGACCTGGAACACAGATGGAAGCTTCT 2100
Db 3070 TCAGCTCTCACCGTATAACCTCAGCACCTCGGAAGACCTGGAACACAGATGGAAGCTTCT 3129
Qy 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAAGTCATGAAGCCACAGGGACTTTTGGTCC 2160
Db 3130 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAAGTCATGAAGCCACAGGGACTTTTGGTCC 3189
Qy 2161 AGCATCTCAGCACCTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
Db 3190 AGCATCTCAGCACCTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3249
Qy 2221 AAACAAAAGTGCCCTACTATATATCAACCAGGACTCAAAACAATTTGCTGGGACCATCCCAA 2280
Db 3250 AAACAAAAGTGCCCTACTATATCAACCAGGACTCAAAACAATTTGCTGGGACCATCCCAA 3309
Qy 2281 AATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAG 2340
Db 3310 AATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAG 3369
Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTTGTGATCTCTTGAGCCTGTC 2400
Db 3370 GACTGCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTTGTGATCTCTTGAGCCTGTC 3429
Qy 2401 A 2401
Db 3430 A 3430

RESULT 13

US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 5339
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-40

Query Match 71.8%; Score 1725; DB 15; Length 5339;

Best Local Similarity 86.1%; Pred. No. 0;				Matches 2068; Conservative 0; Mismatches 0; Indels 333; Gaps 1;			
Qy	1	ACATGCAATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	60				
Db	799	ACATGCAATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA	858				
Qy	61	TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTATCATCACATCACTCTTCCA	120				
Db	859	TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTATCATCACATCACTCTTCCA	918				
Qy	121	AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC	180				
Db	919	AGTTTGCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCC	978				
Qy	181	ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA	240				
Db	979	ACCTAAAGTGACTAAAGAAAGAACATTTTCAGTTACATCATCAAAATGCACATATTCTCAACA	1038				
Qy	241	GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA	300				
Db	1039	GATCAGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTTAAGCCTCGATTCAA	1098				
Qy	301	GAGCTATGCCCTACACACAGSGTGTCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	360				
Db	1099	GAGCTATGCCCTACACACAGSGTGTCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT	1158				
Qy	361	TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	420				
Db	1159	TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG	1218				
Qy	421	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC	480				
Db	1219	TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC	1278				
Qy	481	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	540				
Db	1279	TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338				
Qy	541	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGCAAGCCATCAGGGCCGGTTGG	600				
Db	1339	CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGCAAGCCATCAGGGCCGGTTGG	1398				
Qy	601	TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA	660				
Db	1399	TAAATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA	1458				
Qy	661	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	720				
Db	1459	AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	1518				
Qy	721	TAGCATGGAATAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAACT	780				
Db	1519	TAGCATGGAATAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAACT	1545				
Qy	781	GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAACAAAGGAAATGGAGGAAGA	840				
Db	1546	GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGAACAAAGGAAATGGAGGAAGA	1545				
Qy	841	GCCTCTTGACCTGATCTTGAAGACCTAAACCGCCAAGTACAACAACATAAGGTGCTTCA	900				
Db	1546	GCCTCTTGACCTGATCTTGAAGACCTAAACCGCCAAGTACAACAACATAAGGTGCTTCA	1545				
Qy	901	AGAAGATCTAGAACAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT	960				
Db	1546	AGAAGATCTAGAACAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT	1545				
Qy	961	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	1020				
Db	1546	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	1545				
Qy	1021	AGATCGATGGGCAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAC	1080				

Db	1546	-----	1545
Qy	1081	TCATAGATTACTGCAACAGTTCCCTCTGGACCTTGGAAAAGTTTCTTGCCTGGCTTACAGA	1140
Db	1546	TCATAGATTACTGCAACAGTTCCCTCTGGACCTTGGAAAAGTTTCTTGCCTGGCTTACAGA	1605
Qy	1141	AGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGA	1200
Db	1606	AGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGA	1665
Qy	1201	CTCCAAGGGAGTAAAGAGCTGTGATGAAACAATGCGAAGACCTCCAAGGTGAAATTTGAAGC	1260
Db	1666	CTCCAAGGGAGTAAAGAGCTGTGATGAAACAATGCGAAGACCTCCAAGGTGAAATTTGAAGC	1725
Qy	1261	TCACACAGATGTTTTATCAACAACCTGGATGAAAACAGCCAAAATACTCTGAGATCCCTGGA	1320
Db	1726	TCACACAGATGTTTTATCAACAACCTGGATGAAAACAGCCAAAATACTCTGAGATCCCTGGA	1785
Qy	1321	AGTTTCCGATGATGCAGTCTCTGTTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAG	1380
Db	1786	AGTTTCCGATGATGCAGTCTCTGTTTACAAAGACGTTTGGATAAACATGAACCTCAAGTGGAG	1845
Qy	1381	TGAACCTTCGGAATAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTG	1440
Db	1846	TGAACCTTCGGAATAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTG	1905
Qy	1441	GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA	1500
Db	1906	GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA	1965
Qy	1501	ATTAAGCGGCGAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGT	1560
Db	1966	ATTAAGCGGCGAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGT	2025
Qy	1561	ACATAGGCGCTTCAAGAGGGAATTGAAAACCTTAAGAACTTGAAGAACTCTACCAAGGA	1620
Db	2026	ACATAGGCGCTTCAAGAGGGAATTGAAAACCTTAAGAACTTGAAGAACTCTTCTTGA	2085
Qy	1621	GACTGTACGAATATTTCTGCAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGGA	1680
Db	2086	GACTGTACGAATATTTCTGCAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGGA	2145
Qy	1681	GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA	1740
Db	2146	GCCCAGAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCA	2205
Qy	1741	GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG	1800
Db	2206	GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG	2265
Qy	1801	AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA	1860
Db	2266	AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA	2325
Qy	1861	CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT	1920
Db	2326	CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT	2385
Qy	1921	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGGCC	1980
Db	2386	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGGCC	2445
Qy	1981	TCTGAAAGAGAACGTGAGCCACGTCAAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT	2040
Db	2446	TCTGAAAGAGAACGTGAGCCACGTCAAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT	2505
Qy	2041	TCAGCTCTCACCGTATACTCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCT	2100
Db	2506	TCAGCTCTCACCGTATACTCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCT	2565
Qy	2101	GCAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGCACTTTTGGTCC	2160
Db	2566	GCAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGCACTTTTGGTCC	2625

2161 AGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
2626 AGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2685
2221 AAACAAAGTGGCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCA 2280
2686 AAACAAAGTGGCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCA 2745
2281 AATGACAGAGCTCTACCACTTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAG 2340
2746 AATGACAGAGCTCTACCACTTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAG 2805
2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
2806 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2865
2401 A 2401
2866 A 2866
RESULT 14
JS-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-09-845-416-12
Query Match 70.6%; Score 1695; DB 10; Length 3510;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
21 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 659
61 TGTGATACCACTATCCAGATAGAAGTCCATCTTTAATGTACATCACATCACTCTTCCA 120
660 TGTGATACCACTATCCAGATAGAAGTCCATCTTTAATGTACATCACATCACTCTTCCA 719
121 AGTTTGGCTCAACAAGTGAGCAATTTTCAAGTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
720 AGTTTGGCTCAACAAGTGAGCAATTTTCAAGTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779
181 ACCTAAGTGACTAAAGAAGACATTTTCAGTTACATCAATCAAAATGCACTATTCTCAACA 240
780 ACCTAAGTGACTAAAGAAGACATTTTCAGTTACATCAATCAAAATGCACTATTCTCAACA 839
241 GATCAGGTGAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 300
840 GATCAGGTGAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA 899
301 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGAGCCCCATT 360
900 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGAGCCCCATT 959
361 TCCTTCACAGCAATTTGGAAGCTCTGGAAGACAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 420

960 TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTTGATGGAGAG 1019
421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 480
1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 1079
481 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 540
1080 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
541 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 600
1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1199
601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 660
1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 1259
661 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGSGAATGCCTCAGGGTAGC 720
1260 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTTCAAGATGSGAATGCCTCAGGGTAGC 1319
721 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
1320 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
781 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 840
1380 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 1439
841 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 900
1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1499
901 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGGTAGT 960
1500 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGGTAGT 1559
961 TGATGAATCTAGTGGAGATCAAGCACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1020
1560 TGATGAATCTAGTGGAGATCAAGCACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1619
1021 AGATCGATGGGCAACATCTGTAGTGAAGACCGTGGTGGTCTTTTACAAAGACAC 1080
1620 AGATCGATGGGCAACATCTGTAGTGAAGACCGTGGTGGTCTTTTACAAAGACAC 1676
1081 TCATAGATTACTGCAACAGTTTCCCTCGGACCTGGAAAGTTTCTTGGCTTACAGA 1140
1677 1676
1141 AGCTGAAACAACTGCCAATGTCTTACAGGATGTACCCGTAAAGGAAAGGCTCCTAGAAGA 1200
1677 1676
1201 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
1677 1676
1261 TCACACAGATGTTTATCACAACCTGGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGA 1320
1677 1676
1321 AGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 1380
1677 1676
1381 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGSAAAGCCAGTTCTGACCACTG 1440
1677 1691
1441 GAAGCGTCTGCACCTTTCTCTGCAGGAAGTCTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
1692 GAAGCGTCTGCACCTTTCTCTGCAGGAAGTCTCTGGTGTGGCTACAGCTGAAAGATGATGA 1751

QY	1501	ATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT	1560
DB	1752	ATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT	1811
QY	1561	ACATAGGGCCCTTCAAGAGGGGAATTGAAAACCTAAGAAACCTGTAAATCATGAGTACTCTTGA	1620
DB	1812	ACATAGGGCCCTTCAAGAGGGGAATTGAAAACCTAAGAAACCTGTAAATCATGAGTACTCTTGA	1871
QY	1621	GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA	1680
DB	1872	GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA	1931
QY	1681	GCCCAGAGAGCTGCTCTGAGGAGAGAGGCCCAAGATGTCACTCGGCTTCTACGAAAGCA	1740
DB	1932	GCCCAGAGAGCTGCTCTGAGGAGAGAGGCCCAAGATGTCACTCGGCTTCTACGAAAGCA	1991
QY	1741	GGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	1800
DB	1992	GGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG	2051
QY	1801	AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA	1860
DB	2052	AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA	2111
QY	1861	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGGCGATCTCCT	1920
DB	2112	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGGCGATCTCCT	2171
QY	1921	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC	1980
DB	2172	CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC	2231
QY	1981	TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT	2040
DB	2232	TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT	2291
QY	2041	TCAGCTCTCACGTTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCT	2100
DB	2292	TCAGCTCTCACGTTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCT	2351
QY	2101	GCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGACCTTTGGTCC	2160
DB	2352	GCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGACCTTTGGTCC	2411
QY	2161	AGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	2220
DB	2412	AGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC	2471
QY	2221	AAACAAAGTGCCTACTATATCAACCCAGAGACTCAACAACTTGTCTGGGACCATCCCAA	2280
DB	2472	AAACAAAGTGCCTACTATATCAACCCAGAGACTCAACAACTTGTCTGGGACCATCCCAA	2531
QY	2281	AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAATCTCAGCTTATAG	2340
DB	2532	AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAATCTCAGCTTATAG	2591
QY	2341	GACTGCCATGAAACTCCGAAGACTGCGAAGGCCCTTTTGCTTGATCTCTTGAGCCTGTC	2400
DB	2592	GACTGCCATGAAACTCCGAAGACTGCGAAGGCCCTTTTGCTTGATCTCTTGAGCCTGTC	2651
QY	2401	A	2401
DB	2652	A	2652

RESULT 15

US-09-845-416-31

: Sequence 31, Application US/09845416

Publication No. US20030171312A1

: GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: XIAO, XIAO

APPLICANT: AIAO, AIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

```

; TITLE OF INVENTION: THEREOF
;
; FILE REFERENCE: DB1142
;
; CURRENT APPLICATION NUMBER: US/09/845,416
;
; CURRENT FILING DATE: 2001-04-30
;
; PRIOR APPLICATION NUMBER: 60/200,777
;
; PRIOR FILING DATE: 2000-04-28
;
; NUMBER OF SEQ ID NOS: 36
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 31
;
; LENGTH: 4476
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
US-09-845-416-31

```

; TITLE OF INVENTION: THEREOF

QY	841	GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA	900
Db	2196	GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA	2255
QY	901	AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT	960
Db	2256	AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT	2315
QY	961	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	1020
Db	2316	TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	2375
QY	1021	AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAC	1080
Db	2376	AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGA---	2432
QY	1081	TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGA	1140
Db	2433	-----	2432
QY	1141	AGCTGAACAACACTGCCAATGTCTTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGA	1200
Db	2433	-----	2432
QY	1201	CTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC	1260
Db	2433	-----	2432
QY	1261	TCACACAGATGTTTATCACAACTGTGATGAACAAACAGCCAAAATCCTGAGATCCCTGGA	1320
Db	2433	-----	2432
QY	1321	AGGTTCCGATGATGCAGTCTGTTCACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG	1380
Db	2433	-----	2432
QY	1381	TGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTG	1440
Db	2433	-----CAGTTCTGACCAGTG	2447
QY	1441	GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA	1500
Db	2448	GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA	2507
QY	1501	ATTAAGCCGGCAGGCACCTATTGGAGGCGACCTTCCAGCAGTTTCCAGAGCAGAACGATGT	1560
Db	2508	ATTAAGCCGGCAGGCACCTATTGGAGGCGACCTTCCAGCAGTTTCCAGAGCAGAACGATGT	2567
QY	1561	ACATAGGCGCTTCAAGAGGGGAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTGA	1620
Db	2568	ACATAGGCGCTTCAAGAGGGGAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTGA	2627
QY	1621	GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCCAGGA	1680
Db	2628	GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCCAGGA	2687
QY	1681	GCCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCA	1740
Db	2688	GCCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCA	2747
QY	1741	GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTCGACTCCGCTGACTGGGCAGAG	1800
Db	2748	GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTCGACTCCGCTGACTGGGCAGAG	2807
QY	1801	AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA	1860
Db	2808	AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA	2867
QY	1861	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT	1920
Db	2868	CCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT	2927

QY	1921	CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGGGCC	1980
Db	2928	CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGGGCC	2987
QY	1981	TCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTGCGCAGCTTACCACCTTTGGGCAT	2040
Db	2988	TCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTGCGCAGCTTACCACCTTTGGGCAT	3047
QY	2041	TCAGTCTTCACCGTATTAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT	2100
Db	3048	TCAGTCTTCACCGTATTAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT	3107
QY	2101	GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACCTTTGGTCC	2160
Db	3108	GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACCTTTGGTCC	3167
QY	2161	AGCATCTCAGCACCTTTCTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC	2220
Db	3168	AGCATCTCAGCACCTTTCTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC	3227
QY	2221	AAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAAACAATTGCTGGGACCATCCCAA	2280
Db	3228	AAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAAACAATTGCTGGGACCATCCCAA	3287
QY	2281	AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATCTCTCAGCTTATAG	2340
Db	3288	AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATCTCTCAGCTTATAG	3347
QY	2341	GACTGCCCATGAAACTCCGAAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	2400
Db	3348	GACTGCCCATGAAACTCCGAAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC	3407
QY	2401	A 2401	
Db	3408	A 3408	

Search completed: April 5, 2004, 16:33:42
Job time : 622.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model
Run on: April 4, 2004, 16:52:43 ; Search time 4197.89 Seconds
(without alignments)
17079.796 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 acatgcattcaacatgccca.....tggatctcttgagcctgtca 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mus:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1093.8	45.6	3870	11 BC036103	BC036103 Homo sapi
2	994.4	41.4	5691	29 AY399453	AY399453 Homo sapi
3	864.6	36.0	3056	11 AK044536	AK044536 Mus muscu
4	786.6	32.8	5697	29 AY399455	AY399455 Mus muscu

5	607.8	25.3	5676	29 AY399454	AY399454 Pan trogl
6	605.4	25.2	834	12 BI729851	BI729851 603349511
7	593.4	24.7	595	14 CB177816	CB177816 is21C01.X
8	495.2	20.6	750	12 BI730168	BI730168 603349711
9	486.6	20.3	579	9 AL121550	AL121550 DKFZp762L
10	485.8	20.2	728	14 CB228986	CB228986 AGENCOURT
11	459.6	19.1	1047	14 CB850319	CB850319 MRA-0070
12	424.2	17.7	663	12 BM488464	BM488464 pgm2n.pk0
13	392.6	16.4	3753	11 AK081426	AK081426 Mus muscu
14	389	16.2	644	13 BU313510	BU313510 603540290
15	367.4	15.3	2874	29 AY408546	AY408546 Homo sapi
16	354.6	14.8	2874	29 AY408548	AY408548 Mus muscu
17	324.6	13.5	402	14 CB547284	CB547284 AMGNNUC:S
18	292.4	12.2	9915	29 AY407022	AY407022 Homo sapi
19	286.4	11.9	696	12 BJ075057	BJ075057 BJ075057
20	284.8	11.9	3051	11 BC036095	BC036095 Homo sapi
21	272.6	11.4	2874	29 AY408547	AY408547 Pan trogl
22	271.6	11.3	2334	11 BC011062	BC011062 Mus muscu
23	271.6	11.3	9691	29 AY407024	AY407024 Mus muscu
24	270.2	11.3	1122	14 CF109978	CF109978 Shultzomi
25	269.6	11.2	9096	29 AY407023	AY407023 Pan trogl
26	265.8	11.1	784	12 BG212445	BG212445 RST32032
27	260	10.8	826	14 CD656896	CD656896 AGENCOURT
28	257.8	10.7	1541	11 AK034383	AK034383 Mus muscu
29	256.8	10.7	630	9 AL855376	AL855376 AL855376
30	254	10.6	772	12 BI250598	BI250598 602993659
31	250	10.4	502	9 AL602076	AL602076 DKFZp313B
32	246	10.2	835	12 BI553820	BI553820 603190772
33	233.8	9.7	732	13 BU107880	BU107880 603109863
34	230.2	9.6	763	14 CB518960	CB518960 UI-M-GH0-
35	229.2	9.5	778	14 CB524596	CB524596 UI-M-FY0-
36	226	9.4	250	10 BF963618	BF963618 QV2-NN004
37	216.6	9.0	681	13 BU301653	BU301653 603609005
38	206.4	8.6	479	13 BQ304046	BQ304046 QV2-BT063
39	199.2	8.3	697	13 BU456556	BU456556 603771141
40	187.4	7.8	812	14 CF748008	CF748008 UI-M-HE0-
41	182.4	7.6	645	10 AW467977	AW467977 he29g08.x
42	179.6	7.5	466	14 CD549993	CD549993 B0305E01-
43	178.2	7.4	600	12 BI988528	BI988528 4012-24 M
44	174.6	7.3	851	13 BU201022	BU201022 603952191
45	172	7.2	646	12 BI289102	BI289102 UI-R-DK0-

ALIGNMENTS

RESULT 1	BC036103	3870 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (cdna clone IMAGE:5274415), with apparent retained intron.				
DEFINITION	BC036103.1	GI:23271310			
ACCESSION	BC036103				
VERSION	BC036103.1				
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 3870)				
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,				

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3870)
Strausberg,R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES
source
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN
Query Match 45.6%; Score 1093.8; DB 11; Length 3870;
Best Local Similarity 95.2%; Pred. No. 3.1e-238;
Matches 1128; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
Db 743 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 802
QY 61 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCATCACTCTTCCA 120
Db 803 TGTGTATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCATCACTCTTCCA 862
QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
Db 863 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 922
QY 181 ACCTAAAGTGACTAAAGAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 240
Db 923 ACCTAAAGTGACTAAAGAGAACATTTTCAGTTACATCATCAAAATGCACTATTCTCAACA 982
QY 241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAGCCTCGATTCAA 300
Db 983 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCTAGCCTCGATTCAA 1042

QY 301 GAGCTATGCCTACACACAGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT 360
Db 1043 GAGCTATGCCTACACACAGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT 1102
QY 361 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 420
Db 1103 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1162
QY 421 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 480
Db 1163 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1222
QY 481 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAGTGGTGAAGA 540
Db 1223 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAGTGGTGAAGA 1282
QY 541 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 600
Db 1283 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1342
QY 601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 660
Db 1343 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1402
QY 661 AACTGAAGTACAAGCAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 720
Db 1403 AACTGAAGTACAAGCAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 1462
QY 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 780
Db 1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1522
QY 781 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 840
Db 1523 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1582
QY 841 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACACATAAGGTGCTTCA 900
Db 1583 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACACATAAGGTGCTTCA 1642
QY 901 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 960
Db 1643 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 1702
QY 961 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1020
Db 1703 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1762
QY 1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC 1080
Db 1763 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1822
QY 1081 TCATAGATTACTGCAACAGTTCCCTCTGGACCTGGAAAGATTTCTTGCCTGGCTTACAGA 1140
Db 1823 CCTTCTCAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 1882
QY 1141 AGCTGAAACAACATGCCAATGTCTACAGGATGCTACCCGTAAGGA 1185
Db 1883 AAAAGAAGATGCAGTGAACAAGATTACACAACACTGGCTTTAAAGA 1927

RESULT 2
AY399453
LOCUS
DEFINITION Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY399453 5691 bp DNA linear GSS 12-DEC-2003
genomic survey sequence.
ACCESSION AY399453
VERSION AY399453.1 GI:39755442
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 5691)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 5691)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..5691 <code>/organism=Homo sapiens"</code> <code>/mol_type="genomic DNA"</code> <code>/db_xref="taxon:9606"</code> <1..>5691 <code>/locus_tag="HCM0229"</code>
gene	
ORIGIN	
	Query Match 41.4%; Score 994.4; DB 29; Length 5691; Best Local Similarity 95.0%; Pred. No. 1.Se-215; Matches 1041; Conservative 0; Mismatches 46; Indels 9; Gaps 14
QY	20 AGATATCAATTAGGCATAGAGAAACTACTCGATCTCCTGGAAGATGTTGATACACCTATCCA
Dd	13 AGAACCAAAAAATTGATGTGTAGTGTTTAATGTGCTTACAGATGTTGATACCACCTATCCA
QY	80 GATAAGAAGTCATCTTAATGTACATCACATCACTCTTCCAAAGTTTTGCCCTCAAACAAGTG
Dd	73 GATAAGAAGTCATCTTAATGTACATCACATCACTCTTCCAAAGTTTTGCCCTCAAACAAGTG
QY	140 AGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCCACCTAAAAGTGACTAAAGAA
Dd	133 AGCATTTGAAGCCATCCAGGAAGTGGAATGTTGCCAAGGCCACCTAAAAGTGACTAAAGAA
QY	200 GAACATTTTCAGTTACATCATCAAAATGCACATATTTCTCAACAGATCACGGTCAGTCTAGCA
Dd	193 GAACATTTTCAGTTACATCATCAAAATGCACATATTTCTCAACAGATCACGGTCAGTCTAGCA
QY	260 CAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAG
Dd	253 CAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAG
QY	320 GCTGCTTATGTCACCACTCTGNCCCTACACGGAGCCCCATTTCTTTTCACAGCATTTGGAA
Dd	313 GCTGCTTATGTCACCACTCTGNCCCTACACGGAGCCCCATTTCTTTTCACAGC-----
QY	380 GCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGT
Dd	365 -TCCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGT
QY	440 TATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGGAA
Dd	424 TATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGGAA
QY	500 GCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGACCAGTCTTCTACTCATGAG
Dd	484 GCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGACCAGTCTTCTACTCATGAG
QY	560 GGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGGTAAATATTCTACAATTGGGA
Dd	544 GGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGGTAAATATTCTACAATTGGGA

QY	620	AGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG	679
Db	604	AGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG	663
QY	680	ATGAATCTCTATAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAACAAAAACAAGC	739
Db	664	ATGAATCTCTATAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAACAAAAACAAGC	723
QY	740	AATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGG	799
Db	724	AATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGG	783
QY	800	CTAACAAAAACAGAAGAAAGAACAAAGGAAAAATGGAGGAAGAGCCCTCTTGACCTGATCTT	859
Db	784	CTAACAAAAACAGAAGAAAGAACAAAGGAAAAATGGAGGAAGAGCCCTCTTGACCTGATCTT	843
QY	860	GAAGACCTTAAAAACGCCAAGTACAACAACATTAAGGTGCTTCAGAAGAGATCTAGAACAAGAA	919
Db	844	GAAGACCTTAAAAACGCCAAGTACAACAACATTAAGGTGCTTCAGAAGAGATCTAGAACAAGAA	903
QY	920	CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAGTGGAGAT	979
Db	904	CAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAGTGGAGAT	963
QY	980	CACGCACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC	1039
Db	964	CACGCACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC	1023
QY	1040	TGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACACTCATAGATTACTGCAACAG	1099
Db	1024	TGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACACTCATAGATTACTGCAACAGT	1083
QY	1100	TTCCCCCTGGACCTGG	1115
Db	1084	CTTACTGAAGAACAGG	1099
RESULT 3	AK044536	LOCUS	
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930019F21 product:dystrophin, muscular dystrophy, full insert sequence.		
ACCESSION	AK044536	VERSION	GI:26090404
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 PUBMED 10349636		
REFERENCE	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 MEDLINE 11042159		
REFERENCE	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--38format		

ETIT

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3056)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://location/Qualifiers

FEATURES
source

1.3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930019F21"
/db_xref="MGI:2409854"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"

misc_feature

1.3056
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)"

ORIGIN

Query Match 36.0%; Score 864.6; DB 11; Length 3056;
Best local Similarity 84.0%; Pred. No. 5.1e-186;
Matches 1001; Conservative 0; Mismatches 184; Indels 6; Gaps 2;

QY 1 ACATGCGATTCAACATCGCCAGATATCAATTAGGCATAGAGAACTACTCGATCCTGAAGA 60
|||||
Db 638 ACATGCGCTCAACATTGCAAAATGCCAGTTAGGCATAGAGAACTTCTTGATCCTGAAGA 697
|||||

QY 61 TGTGTATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCATCTCTTCCA 120
|||||

Db 698 TGTTGCTACCACTTATCCAGACAAGAAGTCCATCTTAATGTACATCATCATCTCTTCA 757
|||||

QY 121 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
|||||

Db 758 AGTTTTCCTCAACAAGTGAGCATTTGAAGCCATTTCAAGGAAGTGGAAATGTTGCCAAGGCC 817
|||||

QY 181 ACCT--AAAGTGACTAAAGAAGAACATTTTCAGTTTACATCATCAATGCACTATTCTCA 237
|||||

Db 818 ATCTTCAAAAGTAAGTAGAGAAGAACATTTTCAATTACATCACCAGATGCATTACTCTCA 877
|||||

QY 238 ACAGATCACGGTCAGTCTAGCACAGGGGATATGAGAGAACTTCT---TCCCCTAAGCCTCG 294
|||||

Db 878 ACAGATCACAGTCACTAGCACAGGGCTATGAACAACTTCTTCTCATCTCTTCAAGCCTCG 937
|||||

QY 295 ATTCAAGAGCTATGCCCTACACACAGGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAG 354
|||||

Db 938 ATTCAAGAGTTATGCCCTTACACACAGGGCTGCTTATGTGCCACCTCTGATTCACACAGAG 997
|||||

QY 355 CCCATTTCTTCCAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGAT 414
|||||

Db 998 CCCCTATCTTCCAGCATTTTGAAGCTCCTCCAGAGACAAGTCACTTTCACAGTTTCATTGAT 1057
|||||

QY 415 GGAGAGTGAAGTAAACCTTGGACCGTTTATCAACAGCTTTTAGAAGAAGTATTATCGTGGCT 474
|||||

Db 1058 GGAGACGGAAGTAAATCTGATAGTTTACCAACTGCTTTAGAGAAGTACTTTCATGGCT 1117
|||||

QY 475 TCTTTCTGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGT 534
|||||

Db 1118 TCTTTCTGCCGAGGATACATTTGCGAGCACAGGAGAGATTTTCAATGATGTGGAAGAAGT 1177
|||||

QY 535 GAAAGACCAAGTTTTCATCTAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCG 594
|||||

Db 1178 GAAAGAACAGTTTTCATGCTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACT 1237
|||||

QY 595 GGTGGTAAATTTCTACAAATTTGGGAAGTAAGTGAATTTGGAACAGGAAATTTATCAGAAGA 654
|||||

Db 1238 TGTGGTAAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAATTTATCAGAAGA 1297
|||||

QY 655 TGAAGAACTGAAGTACAAGACAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCTCAG 714
|||||

Db 1298 TGAAGAAGCTGAAGTGAAGAACAAATGAATCTCTCTAAATTTCAAGATGGGAATGTCTCAG 1357
|||||

QY 715 GGTAGTAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTTTAATGGATCTCCAGAAATCA 774
|||||

Db 1358 GGTAGTAGCATGGAAAAACAAAGCAAAATTTACAAAGTTCTAATGGATCTCCAGAAATCA 1417
|||||

QY 775 GAAACTGAAGAAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAATGGA 834
|||||

Db 1418 GAAATTAAGAAGAACTAGATGACTGGTTTAAACAAAAAACTGAAGAGAGAACTAAGAAAATGGA 1477
|||||

QY 835 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAAGGT 894
|||||

Db 1478 GGAAGAGCCTCTTGGACCTGATCTTGAAGATCTTAAATGCTTAAATGCTTAAATGCTTAAAGGT 1537
|||||

QY 895 GCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGT 954
|||||

Db 1538 GCTTCAAGAAGATCTAGAACAGGAGCAGGTCAAGGTCAACTCGCTCACTCACATGGTAGT 1597
|||||

QY 955 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGT 1014
|||||

Db 1598 AGTGGTTGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGGAAAGAACAACTTAAGGT 1657
|||||

QY 1015 ATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTGGTCTTTTACA 1074
|||||

Db 1658 ACTGGGAGATCGATGGGCAAAATATCTGCAGATGGAGTGAAGACCGCTGGATGTTTACA 1717
|||||

QY 1075 AGACACTCATAGATTACTGCAACAGTTTCCCCCTCGACCTGGAAAAGTTTCTGCTGGCT 1134
|||||

Db 1718 AGATATTCTTAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTACATGGCT 1777
|||||

QY 1135 TACAGAAGCTGAACAACTGCGCAATGTCTCTACAGGATGCTACCCGTAAGGA 1185
|||||

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..5676

gene /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>5676 /locus_tag="HCM0229"

ORIGIN

Query Match 25.3%; Score 607.8; DB 29; Length 5676;

Best Local Similarity 60.8%; Pred. No. 2e-127;

Matches 646; Conservative 0; Mismatches 408; Indels 9; Gaps 1;

QY 53 CCTGAAGATGTTGATACCACTATCCAGATAGAAGTCCATCTTAATGTACATCACATCA 112

Db 31 CTTACAGATGTTGCTACCACTATCCAGATAGAAGTCCATTTAATGTACATCACATCA 90

QY 113 CTCTCCAAAGTTTTCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAATGTTG 172

Db 91 CTCTCCAAAGTTTTCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAATGTTG 150

QY 173 CCAAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAAATGCACATAT 232

Db 151 CCCAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCAATTACATCAAAATGCACATAT 210

QY 233 TCTCAACAGATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCT 292

Db 211 TCTCAACAGNN 270

QY 293 CGATTCAAGAGCTATGCGCTACACACAGGCTGCTTATGTCCACCCTCTGACCCTACACGG 352

Db 271 NNN 330

QY 353 AGCCCATTTCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTG 412

Db 331 NNN 390

QY 413 ATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCTGTGG 472

Db 391 NNN 441

QY 473 CTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTG 532

Db 442 NNN 501

QY 533 GTGAAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGC 592

Db 502 NNN 561

QY 593 CGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTATCAGAA 652

Db 562 CGGGTTGGTAATANNNNNNANTTGGGAAGTNNNTGATTGGAAACAGGAAAATTATCAGAA 621

QY 653 GATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTC 712

Db 622 GATGAAGAAACTGAAGTACAAGAGCAGATGAATCTNNNNNATTCAAAGATGGGANNNNNN 681

QY 713 AGGGTAGCTAGCATGGAAGAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCGAAT 772

Db 682 NNN 741

QY 773 CAGAAAACCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAAGGAAAATG 832

Db 742 CAGAAAACCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAAGGAAAATG 801

QY 833 GAGGAAGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACATAAG 892

Db 802 GAGGAAGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACATAAG 861

QY 893 GTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTG 952

Db 862 NNNNNNCAAGAAGATTTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTG 921

QY 953 GTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAACAACTTAAG 1012

Db 922 GTGGTAGTTGATGAANNNNANTGGAGACCCAGCAACTGCTGCTTTGGAAAGAACAACTTAAG 981

QY 1013 GTATTGGGAGATCGATGGGCAACATCTGTAGTAGGACAGAACCCGCTGGGTTCTTTTA 1072

Db 982 GTATTGGGAGATCGATGGGCAACATCTGTAGTAGGACAGAACCCCGCTGGGTTCTTTTA 1041

QY 1073 CAAGACACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGG 1115

Db 1042 CAAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGG 1084

RESULT 6

BI729851

LOCUS 834 bp mRNA linear EST 20-SEP-2001

DEFINITION 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5', mRNA sequence.

ACCESSION BI729851

VERSION BI729851.1 GI:15706864

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 834)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11908 row: e column: 03
High quality sequence stop: 796.

FEATURES Location/Qualifiers

source 1..834

organism="Mus musculus"

mol_type="mRNA"

db_xref="taxon:10090"

clone="IMAGE:5357162"

tissue_type="retina"

lab_host="DH10B (phage-resistant)"

clone_lib="NIH_MGC_94"

note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 25.2%; Score 605.4; DB 12; Length 834;

Best Local Similarity 88.6%; Pred. No. 4.4e-127;

Matches 690; Conservative 0; Mismatches 86; Indels 3; Gaps 3;

QY 1028 TGGGCAACATCTGTAGTGGACAGAGACCGCTGGGTTCTTTTACAGACACTCATAGA 1087

Db 31 TGGGCAACATTCATAAAGAGTAGTGAGCAAGAGGCTGCTTTGGAAAGAACTCATAGA 90

QY 1088 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAA 1147
Db |||||
QY 91 TTACTGCAGCAGTTCCTCTGGACCTGGAGAAGTTTCTTTCTTGATTACCGAAGCAGAA 150
Db |||||
QY 1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGACTCCAAG 1207
Db |||||
QY 151 ACAACTGCCAATGTCCTACAGGACGCTTCCCGTAAGGAGAAGCTCCTAGAAGACTCC-AG 209
Db |||||
QY 1208 GGAGTAAAGAGCTGATGAACAATGGAAGACCTCCAAGGTGAATGAAGCTCACACA 1267
Db |||||
QY 210 GGAGTCAGAGAGCTGATGAACCATGGCAAGATCTCCAAGGAGAAATTGAAACTCACACA 269
Db |||||
QY 1268 GATGTTTATCACACCTGGATGAACACACCCCAAAAAATCCTGAGATCCCTGGAGGTTCC 1327
Db |||||
QY 270 GATATCTATCACAACTCTTGATGAAATGGCCAAAAATCCTGAGATCCCTGGAGGTTCCG 329
Db |||||
QY 1328 GATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAAGTCAAGTGGAGTGAACCTT 1387
Db |||||
QY 330 GATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACCTT 389
Db |||||
QY 1388 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGGAAGCGT 1447
Db |||||
QY 390 CAGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCAAGTTCTGACCAGTGGGAAGCGT 449
Db |||||
QY 1448 CTGCACCTTTCTCTCAGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
Db |||||
QY 450 TTGCATCTTCTCTCAGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGAATGAGC 509
Db |||||
QY 1508 CGGCAGGCACCTATTGGAGGCGACCTTTCAGCAGCTTCAGAAAGCAGATGTATATAGG 1567
Db |||||
QY 510 CGTCAGGCACCCATCGGTGGTGATTTCCAGCAGTTTCAGAAAGCAGAATGATATACATAGG 569
Db |||||
QY 1568 GCCTTCAAGAGGGGAATGAAAACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTA 1627
Db |||||
QY 570 GCCTTCAAGAGGGGAATGAAAACTAAAGAACCTGTAAATCATAGTACTCTTGAGACTGTG 629
Db |||||
QY 1628 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTC-TACCAGGAGCCACG 1686
Db |||||
QY 630 AGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTTACCAGGAGCCACG 689
Db |||||
QY 1687 AGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGA 1746
Db |||||
QY 690 AGAAGTGCCTCCTGGAAGAAAGAGCTCAGAATGTCACTCGGCTCTCTACGAAAGCAGGCTGA 749
Db |||||
QY 1747 GGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTC-CGCTCACTGGCAGAGAAAA 1804
Db |||||
QY 750 AGAGGTCAACGCTGAATGGACCAATTGAACCTGCGCTCAAGCTGATTGGCAGAGAAAA 808
Db |||||

RESULT 7
CB177816/c
LOCUS CB177816 595 bp mRNA linear EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
ACCESSION CB177816
VERSION CB177816.1 GI:28186206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
FEATURES
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 24.7%; Score 593.4; DB 14; Length 595;
Best Local Similarity 99.8%; Pred. No. 2.2e-124;
Matches 594; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 215 CATCATCAAAATGCACCTATCTCAACAGATCACCGTCAGTCTAGCACAGGATATGAGAGA 274
Db |||||
QY 595 CATCATCAAAATGCACCTATCTCAACAGATCACCGTCAGTCTAGCACAGGATATGAGAGA 536
Db |||||
QY 275 ACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACC 334
Db |||||
QY 535 ACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACC 476
Db |||||
QY 335 ACCTCTGACCCCTACACGGAGCCCATTTCCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAG 394
Db |||||
QY 475 ACCTCTGACCCCTACACGGAGCCCATTTCCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAG 416
Db |||||
QY 395 TCATTGGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCCGTTATCAAAACAGCTTTA 454
Db |||||
QY 415 TCATTGGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCCGTTATCAAAACAGCTTTA 356
Db |||||
QY 455 GAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATT 514
Db |||||
QY 355 GAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATT 296
Db |||||
QY 515 TCTAATGATGTGGAAGTGGTGAAGACCCAGTTTCATACACTCATGAGGGGTACATGATGGAT 574
Db |||||
QY 295 TCTAATGATGTGGAAGTGGTGAAGACCCAGTTTCATACACTCATGAGGGGTACATGATGGAT 236
Db |||||
QY 575 TTGACAGCCCATCAGGGCCGGTTGGTAATATTCTACAATTTGGGAAGTAAAGCTGATTGGA 634
Db |||||
QY 235 TTGACAGCCCATCAGGGCCGGTTGGTAATATTCTACAATTTGGGAAGTAAAGCTGATTGGA 176
Db |||||
QY 635 ACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAAT 694
Db |||||
QY 175 ACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAAT 116
Db |||||
QY 695 TCAAGATGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTACATAGATT 754
Db |||||
QY 115 TCAAGATGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTACATAGATT 56
Db |||||
QY 755 TTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAA 809
Db |||||

Db 55 TTAATGGATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAA 1
|||||
BI730168 750 bp mRNA linear EST 20-SEP-2001
603349711F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
mRNA sequence.
BI730168 1 GI:15707181
BI730168 EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 750)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11908 row: f column: 04
High quality sequence stop: 747.
Location/Qualifiers
1..750
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 20.6%; Score 495.2; DB 12; Length 750;
Best Local Similarity 85.5%; Pred. No. 5.7e-102;
Matches 636; Conservative 0; Mismatches 93; Indels 15; Gaps 7;
QY 1000 AGAACAACTTAAGGTATTGGGAGATCGATGGCAACATCTGTAGATGGACAGAACCG 1059
|||||
Db 13 AGAATATCAATACTTCTTGGGGAACATTCACCTAAACATGAGTAAGTGAGCAAGGCTG 72
QY 1060 CTGGGTTCTTTACAGACACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAA 1119
|||||
Db 73 CTTGGGAC-----AGAACTCATAGATTACTGCAGCAGTTCCTCTGGACCTGGAGAA 125
QY 1120 GTTCTTGGCTGG-CTTACAGAGAGCTGAAAC--AACTGCCAATGTCCTACAGGATGCTAC 1176
|||||
Db 126 GTTCTTCTCTGGATTTACGGAAGCAGAAACAGAGCTGCCAATGTCTCTACAGCAGCTTC 185
QY 1177 CCG-TAAGGAAGGCTCCTAGAGACTCCAAGGG--AGTAAAAGAGCTGATGAAACAATG 1233
|||||
Db 186 CCGCTAAGGAGAAAGCTCCTAGCAAGACTCCAGGGGCGAGTCAGAGAGCTGATGAAACCATG 245
QY 1234 GCAAGACCTCAAGGTGAAA-TTGAAGCTCACACAGATGTTTATCATCAACCTGGATGAAA 1292
|||||
Db 246 GCAAGATCTCAAGGAGAAAGTTTGAAACTCACACAGATATCTATCATCAATCTTGATGAAA 305
QY 1293 ACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGAC 1352
|

Db 306 ATGGCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGAAGCACCCCTGTACAAAGAC 365
QY 1353 GTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGT 1412
|||||
Db 366 GTTTGGATAACATGAATTTCAAGTGGAGTGAACCTTCAGAAAAAAGTCTCTCAACATTAGGT 425
QY 1413 CCCATTTGGAAGCCAGTTTCTGACCAGTGGAAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTC 1472
426 CCCATTTGGAAGCAAGTTTCTGACCAGTGGAAAGCGTTTGTCATCTTTCTCTTCAGGAACCTTC 485
QY 1473 TGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACT 1532
486 TTGTTTGGCTACAGCTGAAAAGATGATGAATGAGCCGTGAGGCACCCATCGGTGGTGATT 545
QY 1533 TTCCAGCAGTTCAGAAAGCA-GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACT 1591
546 TCCCAGCAGTTCAGAAAGCACGAATGATATACATAGGCGCTTCAAGAGGGAATTGAAAACT 605
QY 1592 AAAGAACCCTGTAATCATGAGTACTTCTGAGACTGTACGAATATTTCTGACAGAGCAGCCT 1651
606 AAAGAACCCTGTAATCATGAGTACTTCTGAGACTGTGAGAAATATTTCTGACAGAGCAGCCT 665
QY 1652 TTGGAAGGACTAGAGAAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGAGAGAGGCC 1711
666 TTGGAAGGACTAGAGAAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGAGAGAGGCT 725
QY 1712 CAGAATGTCACTCGGCTTCTACGA 1735
|||||
Db 726 CAGAATGTCACTCGGCTCCTACGA 749
RESULT 9
AL121550 579 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp762L078_r1 762 (synonym: hmcl2) Homo sapiens cDNA clone
DEFINITION DKFZp762L078 5', mRNA sequence.
ACCESSION AL121550
VERSION AL121550.1 GI:5927551
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZp762L078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
1..579
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp762L078"
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="762 (synonym: hmcl2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

```
Query Match      20.3%; Score 486.6; DB 9; Length 579;
Best Local Similarity 97.2%; Pred. No. 4.8e-100;
Matches 495; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

2Y 1236 AAGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTATACAAACCTGGATGAAACA 1295
    |||||
Db 71 AGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTATACAAACCTGGATGAAACA 130

2Y 1296 GCCAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGAGTCCCTGTTTACAAAGACGTT 1355
    |||||
Db 131 GCCAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGAGTCCCTGTTTACAAAGACGTT 190

2Y 1356 TGGATAACATGAACATTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCC 1415
    |||||
Db 191 TGGATAACATGAACATTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCC 250

2Y 1416 ATTTGGAAGCCAGTCTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG 1475
    |||||
Db 251 ATTTGGAAGCCAGTCTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG 310

2Y 1476 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 1535
    |||||
Db 311 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 370

2Y 1536 CAGCAGTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGGAAATTGAATAACTAAAG 1595
    |||||
Db 371 CAGCAGTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGGAAATTGAATAACTAAAG 430

2Y 1596 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGG 1655
    |||||
Db 431 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGG 490

2Y 1656 AAGGACTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGCCCCAGA 1715
    |||||
Db 491 AAGGGCTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGCCCCCAG 550

2Y 1716 ATGTCACCTCGGCTTCTACGAAAGCAGGCT 1744
    |||||
Db 551 AATGTCACCTCGCTTCTACGAAAGCAGGCT 579
```

```
RESULT 10
CB228986
LOCUS
DEFINITION
AGENCOURT 11499247 NICHD Rh Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
CB228986
CB228986.1 GI:28280564
EST.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheinae; Macaca.
1 (bases 1 to 728)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers
1..728
/organism="Macaca mulatta"
```

```
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."
```

ORIGIN

```
Query Match      20.2%; Score 485.8; DB 14; Length 728;
Best Local Similarity 92.4%; Pred. No. 7.8e-100;
Matches 511; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 595 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGA 654
    |||||
Db 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAAACAGGAAAATTATCAGAAGA 60

QY 655 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAG 714
    |||||
Db 61 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAG 120

QY 715 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCA 774
    |||||
Db 121 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCA 180

QY 775 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAAAATGGA 834
    |||||
Db 181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAAGAAACAAGGAAAAATGGA 240

QY 835 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAAGGT 894
    |||||
Db 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAAGGT 300

QY 895 GCTTCAAGAAGATCTAGAACAAGAACCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 954
    |||||
Db 301 GCTTCAAGAAGATCTAGAACAAGAACCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 360

QY 955 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTAAGGT 1014
    |||||
Db 361 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAGCAACTAAGGT 420

QY 1015 ATTGGGAGATCGATGGSCAAACATCTGTAGATGGACAGAGCCGCTGGGTCTCTTTTACA 1074
    |||||
Db 421 ATTGGGAGATCGATGGSCAAACATCTGTAGATGGACAGAGCCGCTGGGTCTCTTTTACA 480

QY 1075 AGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCT 1134
    |||||
Db 481 AGACATCCTTCTCAATGGCAACGCTTCTTACTGAAGAACAGTGCCTTTTAGTGATGGCT 540

QY 1135 TACAGAAGCTGAA 1147
    |||||
Db 541 TTCAGAAAAAGAA 553
```

```
RESULT 11
CB850319
LOCUS
DEFINITION
MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
CB850319
CB850319.1 GI:34380806
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1047)
Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
Swaroop,A.
Annotation and analysis of 10,000 expressed sequence tags from
```


LQELRLRLKQHEEAEEAPTLVEGSAEATPDHRNEELLAEARILRQHKSRLETRMQIL
EDHNKQLESOLRLRELLLPSPSESDNGSAGSSSLASSPROEGSHPREKGQTPDTE
VADDVGSKSQDVSLCLEDIMEKLRHAFPSVRSSDVTANTLLAS"

polyA_signal 3731..3736
/note="putative"
polyA_site 3753
/note="putative"

ORIGIN

Query Match 16.4%; Score 392.6; DB 11; Length 3753;
Best Local Similarity 61.3%; Pred. No. 2e-78;
Matches 632; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY	1363	CATGAAC	TTCAAGTGGAGTGAAC	TTCCGGA	AAAGTCTCTCAAC	ATTAGTGCC	CATTGGGA	1422
Db	414	CATGAAT	CTGTGTTGGAATGA	ATAAAGAAAGTCTCA	CAACCTCCGCGCT	CGCCTAGA	473	
QY	1423	AGCCAGT	TTCTGACCACTGGAAGCGT	CTGCACCTTTCTCT	TCGAGSAACTTCTGGTGTG	SGCT	1482	
Db	474	GGCCTT	CTCAGACCTCAGTGGAA	AAACTTCAGCTCCGCT	CCGAGAGATTATTGACTG	SGCT	533	
QY	1483	ACAGCTG	AAAGATGATGAATTAAG	CCCGGCAGGCACCTATT	TGGAGSGCAGCTTTCCAG	CAGT	1542	
Db	534	CAGCCAA	AGGATGAGGATTGTCT	CAGCTCAGCTCCCCT	TGCAAGGGGATGTGGCCCT	TGT	593	
QY	1543	TCAGAAG	CAGAACGATGTACATAG	GGCCCTTCAAGAGGGGA	ATTGNAAACTAAAGAA	ACCTGT	1602	
Db	594	ACAACAG	GAGAAAGGAGACACATG	CAGCCCTCATGGAAGA	AGTCAAGTCTAAGGGCCCC	CTA	653	
QY	1603	AATCATG	AGTACTCTTGAGACTGT	ACGAATATTTCTGAC	AGAGCAGCCCTTTTGGAA	AGGACT	1662	
Db	654	CATCTCT	CTGTGCTTGAATCCGCC	CCAGGCTTTCTGTCC	AGCATCCATTGGAAGA	ATT	713	
QY	1663	AGAGAA	ACTCTACCAGAGCCCGAG	AGACTGCGCTCCTGAG	GAGAGAGCCCGAGA	ATGTCAC	1722	
Db	714	AGAGGAG	CTCAATCTGAGAGTAA	AGATACCTCCCCCAG	ACAGCGGATTCAAAA	CCTTAG	773	
QY	1723	TCGGCTT	CTACGAAAGCAGGCTG	AGGAGTGCAATACTG	AGTGGGAAAAAATTGA	ACCTGCA	1782	
Db	774	TCGCTT	TGTATGGAAGCAGGCA	ACAGTGGCCAGTGAG	CTGTGGGAGAAGCTG	ACAGCCCG	833	
QY	1783	CTCCGCT	GACTGGCAGAAAAAAT	AGATGAGACCCCTTGA	AAGACTCCAGGAACTT	CAAGA	1842	
Db	834	CTGTGTA	GATCAGCATCGCCAC	ATTGAGCATCTCTG	GAACATCTATTGGAG	ATCCAAG	893	
QY	1843	GGCCACG	GATGAGCTGGACCTCA	AGCTGCGCCAAAGCTG	AGGTGATCAAGGGAT	CTCGGA	1902	
Db	894	GGCAATG	GAGGAATTGAGCAGT	ACTTTGACCCAAAGC	AGAGGGAGTCCGAG	CCACGTGGGA	953	
QY	1903	GCCGTG	GGCGATCTCCTCAT	TGACTCTCTCCAAG	ATCACCTCGAGAAAAGT	CAAGGC	1962	
Db	954	GCCCAT	AGGAGATCTCTTTAT	TGATTCCCTCCAG	AGCATATCCAAGCC	ATCAAGCTATT	1013	
QY	1963	TCGAGGA	GAATTGCGCCTCTGAA	AGAGAACGTTGAGCC	ACGTCAATGACCTTGT	CTGCCA	2022	
Db	1014	CAAGAAG	AAATTTCTCTCTGT	GAAAGATGGGGTGA	AGTTAGTGAATGATCT	TGGCCCCA	1073	
QY	2023	GCTTACC	ACTTTGGGCATT	CAGCTCTCACCGT	ATAACCTCAGCACT	CTTGAAGACCTG	2082	
Db	1074	GCTTGCC	ATCTCTGATGTGCACT	TGTCAATGGAGA	ATTCAGGGCTCTGGA	ACAGATCAA	1133	
QY	2083	CACCAG	ATGGAAGCTTTG	CAGGTGGCGT	CGAGCACCGAGT	CAGGCAGCTGCATGA	2142	
Db	1134	CATCCG	TGGAACAGCTCC	AGGTGTCACTGGCT	GAGAGGCTTAAGCA	ACTCCAGGATGC	1193	
QY	2143	CCACAG	GGGACTTTGGTCC	AGCATCTCAGCACT	TTTCCAGCTCTGTCC	AGGGTCCCTG	2202	
Db	1194	CCACCG	GGACTTTGGG	CCCTGGTCA	CAGCACTTCTCTCC	ACTTCTGTCCA	1253	
QY	2203	GGAGAG	AGCCATCTCGCC	AAACAAAGTGC	CTTACTATATCA	CCAGAGACTCAA	2262	
Db	1254	GGAAG	AGGCAATTTCTCC	CAATAAAGTCCC	CTACTACATCA	CCACAGGGCTCAG	1313	

QY	2263	TTGCTGG	ACCACATCC	CAAAATG	ACAGAGCT	CTACCA	GTCTTAGCT	GACCTGA	2322
Db	1314	ATGCTGG	ACCACATC	CTAAGAT	GACTGAGT	TATAC	CAAAACCT	TAGCCG	1373
QY	2323	CAGATT	CTCAGCTT	ATAGGACT	GCATGAA	AACTCC	GAAAGCT	GCAGAA	2382
Db	1374	TAAGTT	CTCAGCTT	TACCGC	ACTGCC	ATGAAG	CTCCG	CAGAGT	1433
QY	2383	GGATCT	CTCTGA	2393					
Db	1434	GGATCT	GGTAA	1444					

RESULT 14
BU313510
LOCUS
DEFINITION
603540290F1 CSEQCHN61 644 bp mRNA linear EST 28-NOV-2002
sequence.
BU313510
BU313510.1 GI:25821511
EST.
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 644)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..644
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="ChEST508c24"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN61"
/note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 16.2%; Score 389; DB 13; Length 644;
Best Local Similarity 75.5%; Pred. No. 8.7e-78;
Matches 482; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 1425 CCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGGCTAC 1484

Db 7 CGAGCACAGACCAGTGGAGCGTTTACATCTCTCTTTCAGGAACCTTTGGCATGGGTGC 66
QY 1485 AGCTGAAAGATGATGAATTAAGCCGCGGACACCTATTGGAGGCGACTTTCCAGCAGTTC 1544
Db 67 AATTGAAGGAGGATGAATTAACAACAGCAAGCACCCATTGGTGGAGATATCCCACTGTGC 126
QY 1545 AGAAGCAGAACGATGTACATAGGGCCTTTCAGAGGGAATTGAAATACTAAAGAACCTGTAA 1604
Db 127 AGAAGCAGAATGATGTTCATAGGACTTTTCAGAGGGAGCTGAAATACTAAAGAACCTGTAA 186
QY 1605 TCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAG 1664
Db 187 TCATGATGCACCTTGAGACTGTGCGACTCTTCCTGGCAGATCAACCAGTAGAGGGACTGG 246
QY 1665 AGAAACTCTACAGGAGCCCGACAGAGCTGCTCCTTGAGGAGAGAGCCCGAGATGTCACTC 1724
Db 247 AAAAGGTCTATCCAGAAACCAAGAGACCTATCACCTTGAGGAGAGGGCCCGAGATGTCACTA 306
QY 1725 GGCCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACT 1784
Db 307 AAGTTCTCCGAAAGCAGCAGATGATGTCAAACTGAGTGGGATAAGCTAAATCTACGTT 366
QY 1785 CCGCTGACTGGCAGAGAAAAATAGATGAGAGCCCTTGAAAGACTCCAGGAACTTCAAGAGG 1844
Db 367 CTGCTGATTGGCAAAAGAGATAGATGATGCTCTTGAAAGACTCGAGGGTCTTCAGGAGG 426
QY 1845 CCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCAGC 1904
Db 427 CAATGGATGAACCTAGACCTGAAACTGCGCCAGGCTGAAGCATTCAGGGATCCTGGCAGC 486
QY 1905 CCGTGGCGGATCTCCTCAATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTC 1964
Db 487 CAGTGGGGGATCTGCTGATAGACTCTCTGCAGGATCACTTAGAAAAAGTCAAGGTTTATC 546
QY 1965 GAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGC 2024
Db 547 GAGCAGAAATGGTGGCCCTTANAGAGAAAGGTGCATCAAGTCAATGAGCTGGCTCACCGGT 606
QY 2025 TTACCACCTTGGGCAATCAGCTCTCACCGTATACCTC 2062
Db 607 TCGCTCCCCCTGATATTACGTTCTCTCCCATACACTCTC 644

RESULT 15
AY408546
LOCUS
DEFINITION Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408546
VERSION AY408546.1 GI:39764517
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2874)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2874)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2874
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2874
/gene="DRP2"
/locus_tag="HCM3260"
ORIGIN
Query Match 15.3%; Score 367.4; DB 29; Length 2874;
Best Local Similarity 57.5%; Pred. No. 1e-72;
Matches 593; Conservative 0; Mismatches 438; Indels 0; Gaps 0;
QY 1363 CATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGTCCCATTTGGA 1422
Db 234 CATGAATCTGTGTGGATGAAATAAAAAAGAGTCTCAACCTCCGGCTCGCCTAGA 293
QY 1423 AGCCAGTCTTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCT 1482
Db 294 GGCCTTCTCAGACCACAGTGGAAAGCTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCT 353
QY 1483 ACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGGCGACTTTCCAGCAGT 1542
Db 354 CAGCCAAAAGGATGAGGAGTGTTCAGTCTCAGCTGCCCCACAGGGGATGTGGCCCTGGT 413
QY 1543 TCAGAAACAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACATAAAGAACCTGT 1602
Db 414 GCAACAGAGAGAGGAGACACATCGGCCCTTTATGGAAGAAGTCAAGTCTCGGGGCCCTTA 473
QY 1603 AATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACT 1662
Db 474 CATCTATTCTGTGCTGGAGTCAGCTCAGGCCCTTCTGTCCAGCACCCATTGAGGAGTT 533
QY 1663 AGAGAAACTCTACCAGGAGCCCGACAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTAC 1722
Db 534 AGAGGAGCCTCATTTCTGAGAGCAAGATACCTCCCGAAACAGCGGATCCAGAATCTCAG 593
QY 1723 TCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCA 1782
Db 594 CCGCTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACCTGTGGAGAGTTGACAGCCCG 653
QY 1783 CTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGA 1842
Db 654 CTGTGTGACCCAGCACCGTACATTGAGCGGACTCTGGAGCAGCTCTTGGAGATTCAAGG 713
QY 1843 GGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCA 1902
Db 714 GGCAATGGAGGAACTAAGCACTACTCTGAGCCCAAGCTGAGGGAGTCCGAGCCACTTGGGA 773
QY 1903 GCCGTGGCGGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCAC 1962
Db 774 GCCCATTTGGGATCTCTTCATTGATTCACTCCACAGACACATCCAGGCTATTAAAGCTGTT 833
QY 1963 TCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTGCGCA 2022
Db 834 CAAAGAAGAAATTCTCCCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCCACCA 893
QY 2023 GCTTACCACTTTGGGCATTGAGCTCTCACCGTATPAACTCAGCACTCTTGAAGACCTGAA 2082
Db 894 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAAATCCAGGCCCTGGAAACAGATCAA 953
QY 2083 CACCAGATGGAAGCTTCTGAGGTGGCCGTTCGAGGACCGAGTCAGGCAGTGCATGAAGC 2142
Db 954 CGTCCGATGGAACAACACTACAGGCGTCAGTTAGTGAGAGGCTTAAAGCAGCTCCAGGATGC 1013
QY 2143 CCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTG 2202
Db 1014 CCACCGGAGCTTTGGCCCTGGTCCAGCACATTTCTCTCCTCTCTGTCCAGGTTCCCTG 1073
QY 2203 GGAGAGAGCCATCTCGCCCAAAACAAAGTGGCCCTACTATATCAACCAGGAGACTCAAAAC 2262

Db	1074	GGAAAGAGCAATTTCACCCCAATAAAGTTCCTACTACATCAANNNNNNNNNNNNNNNNNN	1133
Qy	2263	TTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGT	2322
Db	1134	NN	1193
Qy	2323	CAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAAGCCCTTTTGCTT	2382
Db	1194	TAAGTTCTCAGCTTATCGCACTGCCATGAAACTCCGAGAGTCCAGAAAGCCCTGCGCTT	1253
Qy	2383	GGATCTCTTGA	2393
Db	1254	GGACCTGGTAA	1264

Search completed: April 5, 2004, 08:15:07
Job time : 4201.22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:58:13 ; Search time 63.0939 Seconds
(without alignments)
11443.127 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgttccaatttg 1301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	100.0	5952	4	US-09-687-875A-1
2	1290	99.2	13977	4	US-09-484-970B-60
C 3	1135.2	87.3	19307	3	US-08-836-022A-10
C 4	1135.2	87.3	19307	3	US-09-427-048A-10
5	553.8	42.6	6045	4	US-09-091-501B-7
6	553.8	42.6	10320	4	US-09-091-501B-9
7	540.2	41.5	3915	4	US-09-976-594-93
C 8	50	3.8	7218	1	US-08-232-463-14
9	44.2	3.4	2574	4	US-09-668-313A-10
10	42.8	3.3	1690	4	US-09-620-312D-69
11	42.8	3.3	7812	3	US-09-368-590-1
12	38.6	3.0	1995	1	US-08-425-069-3
13	38.6	3.0	1995	2	US-08-317-844B-3
14	38.4	3.0	7672	4	US-09-220-132-24
15	38.2	2.9	428	4	US-09-668-313A-3
16	38.2	2.9	4439	4	US-09-668-313A-17
17	34.6	2.7	474	4	US-09-621-976-18033
C 18	34.6	2.7	2277	1	US-08-676-967-5
C 19	34.6	2.7	2277	1	US-08-676-974-5
C 20	34.6	2.7	2277	2	US-09-098-487-5
21	34.4	2.6	1047	4	US-09-671-950-1
22	34.4	2.6	1047	4	US-09-671-950-3
23	34.4	2.6	1047	4	US-09-671-950-5
24	34.4	2.6	1047	4	US-09-671-950-7
25	34.4	2.6	1047	4	US-09-671-950-9
26	34.4	2.6	1047	4	US-09-671-950-11
27	34.4	2.6	1047	4	US-09-671-950-13

28	34.2	2.6	750	4	US-08-961-527-370	Sequence 370, Appl
29	34.2	2.6	1620	4	US-08-858-207A-56	Sequence 56, Appl
30	34.2	2.6	2800	2	US-08-874-138-1	Sequence 1, Appli
31	34.2	2.6	2800	2	US-08-874-138-5	Sequence 5, Appli
32	34.2	2.6	2800	3	US-08-879-941-1	Sequence 1, Appli
33	34.2	2.6	2800	3	US-08-879-941-3	Sequence 3, Appli
34	34.2	2.6	2800	4	US-09-747-116-1	Sequence 1, Appli
35	34.2	2.6	2800	4	US-09-747-116-3	Sequence 3, Appli
36	34.2	2.6	32768	4	US-08-961-527-71	Sequence 71, Appl
37	34	2.6	648	4	US-09-252-991A-10033	Sequence 10033, A
C 38	34	2.6	762	4	US-09-252-991A-9821	Sequence 9821, Ap
39	34	2.6	1644	4	US-09-252-991A-10161	Sequence 10161, A
C 40	34	2.6	2235	3	US-09-153-804-2	Sequence 2, Appli
C 41	33.8	2.6	505	4	US-09-621-976-15639	Sequence 15639, A
42	33.8	2.6	1603	1	US-08-625-209A-1	Sequence 1, Appli
43	33.8	2.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
44	33.8	2.6	3489	4	US-09-298-568-1	Sequence 1, Appli
45	33.8	2.6	3489	4	US-09-410-399-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match	100.0%;	Score 1301;	DB 4;	Length 5952;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1301;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGACTTCCAGCAGTTTCAGAACGATGATACATAGGGCCCTTCAGAGGGAATTGAA	60	
Db	3402	CGACTTCCAGCAGTTTCAGAACGATGATACATAGGGCCCTTCAGAGGGAATTGAA	3461	
QY	61	AACATAAGAACCTGTAATCATGACTCTTTGAGACTGTACGAATATTTCTGACAGACCA	120	
Db	3462	AACATAAGAACCTGTAATCATGACTCTTTGAGACTGTACGAATATTTCTGACAGACCA	3521	
QY	121	GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGCTGCCCTCCTGAGGAGAG	180	
Db	3522	GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGCTGCCCTCCTGAGGAGAG	3581	
QY	181	AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCACTACTAGTGGGA	240	
Db	3582	AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCACTACTAGTGGGA	3641	
QY	241	AAAATTGAACCTGCACCTCCGTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACT	300	

Db 3642 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3701
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db 3702 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3761
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCTCAAGATCACCTCGA 420
Db 3762 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCTCAAGATCACCTCGA 3821
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db 3822 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 3881
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 3882 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 3941
QY 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGCCGAGTCAG 600
Db 3942 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGCCGAGTCAG 4001
QY 601 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCAGSCATCTCAGCACTTCTTTCCACGTC 660
Db 4002 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCAGSCATCTCAGCACTTCTTTCCACGTC 4061
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720
Db 4062 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 4121
QY 721 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 780
Db 4122 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 4181
QY 781 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 4182 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 4241
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTTGACCCAGCA 900
Db 4242 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTTGACCCAGCA 4301
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTTGACCCAC 960
Db 4302 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTTGACCCAC 4361
QY 961 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGG 1020
Db 4362 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGG 4421
QY 1021 TATGTCTTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTTG 1080
Db 4422 TATGTCTTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTTG 4481
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAAGTACAG 1140
Db 4482 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAAGTACAG 4541
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 4542 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4601
QY 1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTTGCATCCTTTGGGGG 1260
Db 4602 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTTGCATCCTTTGGGGG 4661
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 4662 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTTG 4702

RESULT 2

US-09-484-970B-60

; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 99.2%; Score 1290; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 8775
QY 61 AACTAAAGAACCTGTATCATATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGAACCTGTATCATATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 8835
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 180
Db 8836 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 8895
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATFACTGAGTGGGA 240
Db 8896 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATFACTGAGTGGGA 8955
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 8956 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 9015
QY 301 CCAGGAACTTCAAGAGGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 9016 CCAGGAACTTCAAGAGGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 9075
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCTCAAGATCACCTCGA 420
Db 9076 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCTCAAGATCACCTCGA 9135
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAAGAGAACGTTGAGCCACGTCAA 480
Db 9136 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAAGAGAACGTTGAGCCACGTCAA 9195
QY 481 TGACCTTGCTCGCCAGCTTTACCACTTTGGGCATTTGAGCTTCCAGCTCTCACCGTATAACCTCAGCAC 540
Db 9196 TGACCTTGCTCGCCAGCTTTACCACTTTGGGCATTTGAGCTTCCAGCTCTCACCGTATAACCTCAGCAC 9255
QY 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 9256 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 9315
QY 601 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
Db 9316 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 9375
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720

QY 782 GACCTGAATAATGTGATCTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAG 841
Db 5197 GACCTGAATAATGTGATCTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAG 5138
QY 842 AAGGCGCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGATGCTTGGACCAAGCAC 901
Db 5137 AAGGCGCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGATGCTTGGACCAAGCAC 5078
QY 902 AACCTCAAGCAAAATGACCAAGCAAGCAACAAATTTGGTCAACGCTCTCTGCTGGAT 961
Db 5077 AACCTCAAGCAAAATGACCAAGCAAGCAACAAATTTGGTCAACGCTCTCTGCTGGAT 5018
QY 962 ATTTATGACCGCTGGAGCAAGCAAGCAACAAATTTGGTCAACGCTCTCTGCTGGAT 1021
Db 5017 ATTTATGATCGTCTGGAGCAAGCAAGCAACAAATTTGGTCAACGCTCTCTGCTGGAT 4958
QY 1022 ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 1081
Db 4957 ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 4898
QY 1082 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAGA 1141
Db 4897 CTGTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCACACTTGGAGACAAGTACAGA 4838
QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGGAGGCTGGGCTC 1201
Db 4837 TACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGTGACCAGCGGAGGCTGGGCTC 4778
QY 1202 CTTCTGATGATTTCTATCCAAATTTCAAGACAGTGGTGAAGTTGCATCTTTGGGGGC 1261
Db 4777 CTTCTGATGATTTCTATTTCAAAATTTCAAGACAGTGGTGAAGTTGCTTCTTTGGGGGC 4718
QY 1262 AGTAACATGAGCCAAAGTGTCCGAGCTGCTTCCAAATTTG 1301
Db 4717 AGTAACATGAGCCGAGTGTCCGAGCTGCTTCCAAATTTG 4678

RESULT 4
US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Fisher, James M.
Chen, Krishna J.
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10
Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAA 61
Db 5977 GATTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAA 5918
QY 62 ACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 121
Db 5917 ACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTGAGAAATATTTCTGACAGAGCAG 5858
QY 122 CCTTTGGAGAGGACTAGAGAAACTCTACAGGAGCCACAGAGAGCTGCCTCTCTGAGGAGAGA 181
Db 5857 CCTTTGGAGAGGACTAGAGAAACTCTACAGGAGCCACAGAGAACTGCCTCTCTGAGGAAAGA 5798
QY 182 GCCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAA 241
Db 5797 GCTCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAACGCTGATGGGAC 5738
QY 242 AAATTGAACCTGCACCTCCGCTGACTGGCAGAGAGAAAATAGATGAGACCCCTTGAAGACTC 301
Db 5737 AAATTGAACCTGCCTCAGCTGATGGCAGAGAGAAAATAGATGAGACTCTTGAAGACTC 5678
QY 302 CAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATC 361
Db 5677 CAGGAACCTTCAAGAGGCTGCGGATGAACTGGACCTCAAGTTGCGCCCAAGCTGAGGTGATC 5618
QY 362 AAGGATCTCTGGCAGCCCGTGGGCGATCTCCTCATTTGATCTCTCCAAGATCACTCGAG 421
Db 5617 AAGGATCTCTGGCAGCCCGTGGGCGATCTCCTCATTTGATCTCTCCAAGATCACTCGAG 5558
QY 422 AAAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAAACGTGAGCCACGTCAAT 481
Db 5557 AAAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAAATGTCAATCGTGTCAAT 5498
QY 482 GACCTTGTCTGGCAGCTTACCACTTTGGGCACTTCAGCTCTCAGCTTAACTCAGCACT 541
Db 5497 GACCTTGCACATCAGCTGACCACTGGGCACTTCAGCTCTCAGCTTAACTCAGCACT 5438
QY 542 CTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGAGGTGGCGTTCGAGGACCGAGTCAGG 601
Db 5437 TTGGAAGATCTGAATACCAAGATGGAGGCTTCTACAGGTGGCTGTGGAGGCCGTGTGAGA 5378
QY 602 CAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCT 661
Db 5377 CAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTCA 5318
QY 662 GTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCAC 721
Db 5317 GTTCAGGCTCCCTGGGAGAGAGCCATCTCACCAAAACAAAGTGCCTTACTATATCAACCAC 5258
QY 722 GAGACTCAAAACAACTTGTGGGACCACTCCCAAAATGACAGAGCTTACCAGTCTTTAGCT 781
Db 5257 GAGACCCAAACCACTTGTGGGACCACTCCCAAAATGACAGAGCTTACCAGTCTTTAGCT 5198
QY 782 GACCTGAATAATGTGATCTCAGCTTATAGGACTGCCATGAACTCCGAGAGACTGCAG 841
Db 5197 GACCTGAATAATGTGATCTCAGCTTATAGGACTGCCATGAACTCCGAGAGCTGCAG 5138

842 AAGGCCCTTGTGGATCTCTTGGAGCCTGTGAGCCTGTGATGCTGATGCTGGACCGAC 901
|||||
5137 AAGGCCCTTGTGGATCTCTTGGAGCCTGTGAGCCTGTGATGCTGATGCTGGACCGAC 5078
|||||
902 AACCTCAAGCAAAATGACAGCCCATGGATATCTGACAGATTATTAATGTTTGACCACT 961
|||||
5077 AACCTCAAGCAAAATGACAGCCCATGGATATCTGACAGATAATTAACCTGTTGACTACA 5018
|||||
962 ATTTATGACCGCTGGAGCAAGAGACACAACAAATTTGGTCAACGTCCTCTCTGCGTGGAT 1021
|||||
5017 ATTTATGATCGTCTGGAGCAAGAGACACAACAAATCTGGTCAATGTCCCTCTCTGTGTGGAT 4958
|||||
1022 ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTC 1081
|||||
4957 ATGTGCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTC 4898
|||||
1082 CTGCTTTTAAAACTGGCATCAATTTCCCTGTGTAAGACACATTTGGAAAGACAAGTACAGA 1141
|||||
4897 CTGCTTTTAAAACTGGCATCAATTTCTGTGTAAAGCACACATTTGGAAAGACAAGTACAGA 4838
|||||
1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGCCCTC 1201
|||||
4837 TACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAAGCGTAGGCTGGGTCTT 4778
|||||
1202 CTCTGCTGATTTCTATCCAAATTCGAAGACAGTGGGTGAAGTTGCAATCCTTTGGGGGC 1261
|||||
4777 CTCTGCTGATTTCTATTCAAAATCCCAAGACAGTTGGGTGAAGTTGCTTCTCTTTGGGGGC 4718
|||||
1262 AGTAACATTTAGCCCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
|||||
4717 AGTAACATTTAGCCCGAGTGTGAGGAGCTGCTTCCAAATTTG 4678
|||||

RESULT 5
IS-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open

IS-09-091-501B-7

Query Match 42.6%; Score 553.8; DB 4; Length 6045;
Best Local Similarity 64.7%; Pred.No. 1.1e-169;
Matches 844; Conservative 0; Mismatches 452; Indels 9; Gaps 1;

QY 2 GACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAA 61
|||
Db 3506 GATGTTCCAGCCTTACAGCTCCAGTATGACCAATTGTAAAGCCCTGAGACGGGAGTTAAAG 3565
|||
QY 62 ACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 121
|||
Db 3566 GAGAAAGAAATATTTCTGCTCTGAAATGCTGCGACCGGCGGAGTTTCTTGGCTGATCAG 3625
|||
QY 122 CCTTTGGAGAGACTAGAGAAACT-----CTACAGGAGCCCCAGAGAGCTGCCCTCCT 172
|||
Db 3626 CCAATTGAGGCCCTGAAGAGCCCAAGAAACCTACAAATCAAAAACAGAAATTAACCTCCT 3685
|||
QY 173 GAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 232
|||
Db 3686 GAGGAGAGAGCCCGAAAGATTGCCAAAGCCATGGCGCAAAACAGTCTTCTGAAGTCAAAGAA 3745
|||
QY 233 GAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 292
|||
Db 3746 AAATGGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGCAAGTGGACAAGGCATTG 3805
|||
QY 293 GAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 352
|||
Db 3806 GAGAAACTCAGAGACCTGCAGGGAGCTATGGATGACCTGGACGCTGCATGAAGAGGCA 3865
|||
QY 353 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGAT 412
|||
Db 3856 GAGTCCGTGCGGAATGGCTGGAAGCCCGTGGGAGACTTACTCATTGACTCGCTGCAGGAT 3925
|||
QY 413 CACCTCGAGAAAGTCAAGGCACCTTGAGAGAGAAATTTGGGCTCTGAAAGAGAACGCTGAGC 472
|||
Db 3926 CACATTGAAAAAATCATGGCATTTAGAGAAAGAAATTGACCAAAATCAACTTTAAAGTTAAA 3985
|||
QY 473 CACGTCAATGACCTTGTCTGCCAGCTTACCACATTTGGGCAATTCAGCTCTCACCGTATAAC 532
|||
Db 3986 ACGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTGACCTGCATCCCTCTCTAAAGATG 4045
|||
QY 533 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTGGCGTCGAGGAC 592
|||
Db 4046 TCTCGCCAGCTAGATGACCTTAATATGCGATGGAACCTTTTACAGGTTTCTGTGGATGAT 4105
|||
QY 593 CGAGTCAGGCGAGCTGCATGAAGCCCAAGGAGGACTTTGGTCAGCATCTCAGCACTTTCTT 652
|||
Db 4106 CGCCTTAAACAGCTTCAGGAAGCCCAAGAGATTTTGGACCATCTCTCAGCATTTTCTC 4165
|||
QY 653 TCCACGCTCTGCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTACTAT 712
|||
Db 4166 TCTACGTCAGTCCAGCTGCCGTGGCAAGATCCATTTTACATAATAAAGTCCCTATTAC 4225
|||
QY 713 ATCAACCACGAGACTCAAAACAACCTTCTGTTGGACCATCCCAAAATGACAGAGCTCTACCAG 772
|||
Db 4226 ATCAACCATCAAAACACAGACCCTGTGGGACCATCTCTAAAATGACCGAATCTTTCAA 4285
|||
QY 773 TCTTTAGCTGACCTGAATAATGTGATCTCTCAGCTTATAGGACTGCCATGAAACTCCGA 832
|||
Db 4286 TCCCTTGTGACCTGAATAATGTACGTTTTTCTGCCCTACCGTACAGCAATCAAAATCCGA 4345
|||
QY 833 AGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTG 892
|||
Db 4346 AGACTACAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACAAATGAAATTTTC 4405
|||
QY 893 GACCAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCTCTGCAGATTATTAAATGT 952
|||
Db 4406 AAACAGCACAAAGTTGAACCAAAATGACCAGCTCCTCAGTGTCCAGATGTCACTCACTGT 4465
|||
QY 953 TTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTC 1012
|||
Db 4466 CTGACAACAACCTTATGATGGACTTGAGCAAAATGCATAAGGACCTGGTCAACGTTCCACTC 4525
|||
QY 1013 TGGGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGG 1072
|||
Db 4526 TGTGTTGATATGTGTCTCAATTGGTTGCTCAATGTTGTTGCTCAATGTCTATGACACGGGTGCGAACTGGAAAA 4585
|||

QY 1073 ATCCGTGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 1132
|||
Db 4586 ATTAGAGTGCAGAGTCTGAAGATTGGATTAAATGCTCTCTCCAAAGGTCTCTTTGGAAGAA 4645
|||
QY 1133 AAGTACAGATACCTTTTCAAGCAAGTGCACAGTTTCAACAGGATTTTGTGACCAAGCGCAGG 1192
|||
Db 4646 AAATACAGATATCTCTTAAGGAAGTTGGGGCCGACAGAAATGTGTGACCAAGGCAG 4705
|||
QY 1193 CTGGGCTCTCTTGTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTGCATCC 1252
|||
Db 4706 CTGGGCTGTACTTTCATGATGCCATCCAGATCCCCCGGCAGTAGGTGAAGTAGCAGCT 4765
|||
QY 1253 TTTGGGGCAGTAACATTTAGCCAAAGTCTCCGGAGCTGCTTTCCAA 1297
|||
Db 4766 TTTGGAGGCAGTAATATTGAGCCTAGTGTTCGCAGCTGCTTTCCAA 4810
|||

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open

US-09-091-501B-9
Query Match 42.6%; Score 553.8; DB 4; Length 10320;
Best Local Similarity 64.7%; Pred. No. 1.5e-169;
Matches 844; Conservative 0; Mismatches 452; Indels 9; Gaps 1;
QY 2 GACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAA 61
|||
Db 7781 GATGTTCCAGCCTTACAGCTCCAGTATGACCATTTGAAGCCCTGAGACGGGAGTTAAAG 7840
|||
QY 62 ACTAAGAACCCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAG 121
|||
Db 7841 GAGAAAGAAATATTCTGTCCTGAATGCTGTCGACCGCCCGAGTTTCTTGGCTGATCAG 7900
|||
QY 122 CCTTTGGAAGGACTAGAAACT-----CTACCAGGAGCCACAGAGCTGCCTCCT 172
|||
Db 7901 CCAATTGAGGCCCTTGAGAGCCCAAGAGAACCTTACAATCAAAAACAGAATTAACCTCT 7960
|||
QY 173 GAGGAGAGAGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 232
|||
Db 7961 GAGGAGAGAGCCCAAGATTTGCCAAAGCCATGGCAACACAGTCTTCTGAAGTCAAGAA 8020
|||

QY 233 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 292
|||
Db 8021 AAATGGGAAAGTCTAAATGCTGTAACCTAGCAATTTGGCAAAAGCAAGTGACAAGGCATTG 8080
|||
QY 293 GAAAGACTCCAGGACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCT 352
|||
Db 8081 GAGAAACTCAGAGACCTGCAGGGAGCTATGGATGACCTGGACGCTGACATGAAGGAGGCA 8140
|||
QY 353 GAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGAT 412
|||
Db 8141 GAGTCCGTGCGGAATGGCTGGAAAGCCCCGTGGGAGACTTACTCATTTGACTCGCTGCAGGAT 8200
|||
QY 413 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 472
|||
Db 8201 CACATTGAAAAAATCATGGCATTTAGAGAGAGAAATTGCACCAATCAACTTTAAAGTTAAA 8260
|||
QY 473 CACGTCAATGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC 532
|||
Db 8261 ACGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTGACCTGCATCCCTCTCTAAAGATG 8320
|||
QY 533 CTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTTGGCCGTCGAGGAC 592
|||
Db 8321 TCTGCCAGCTAGATGACCTTAATATGCGATGGAAACTTTTACAGGTTTCTGTGGATGAT 8380
|||
QY 593 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTT 652
|||
Db 8381 CGCCTTAAACAGCTTCAGGAAGCCCACAGAGATTTTGGACCATCTCTCTCAGCATTTTCTC 8440
|||
QY 653 TCCACGTCTGTCCAGGTCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTAT 712
|||
Db 8441 TCTACGTCACTCCAGCTGCCGTGGCAAGATCCATTTCACATAATAAAGTGCCCTATTAC 8500
|||
QY 713 ATCAACCACGAGACTCAAAACAACTTGTCTGGACCATCTCCAAATTCACAGAGCTCTACCAG 772
|||
Db 8501 ATCAACCATCAAAACACAGACCCACCTGTTGGACCATCTTAAATGACCGAACTCTTTCAA 8560
|||
QY 773 TCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGA 832
|||
Db 8561 TCCCTTGCTGACCTGAATAATGTACGTTTCTGCTCCTACCGTACAGCAATCAAAATCCGA 8620
|||
QY 833 AGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCCCTTG 892
|||
Db 8621 AGACTACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACAAACAATGAAATTTTC 8680
|||
QY 893 GACCAGCACAACTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTGT 952
|||
Db 8681 AAACAGCACAAAGTTGAACCAAAATGACCAAGCTCCTCAGTGTTCAGATGTCACTCAACTGT 8740
|||
QY 953 TTGACCACTATTATGACCGCCTGGAGCAAGACAGACAAACAATTTGGTCAACGTCCCTCTC 1012
|||
Db 8741 CTGACAACAACTTATGATGGACTTGAGCAAAATGACATAAGGACCTGGTCAACGTTCCTACTC 8800
|||
QY 1013 TGCCTGGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGG 1072
|||
Db 8801 TGTGTTGATATGTGCTCAATTTGGTTGCTCAATGTCTATGACACGGGTGGAACCTGGAATA 8860
|||
QY 1073 ATCCGTCTCCTGCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 1132
|||
Db 8861 ATTAGAGTGCAGAGTCTGAAGATTGGATTAAATGTCTCTCTCCAAAGGTCTCTTGAAGAA 8920
|||
QY 1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGG 1192
|||
Db 8921 AAATACAGATATCTCTTTAAGGAAGTTGCGGGSCCGACAGAAATGTGTGACCAGAGGCAG 8980
|||
QY 1193 CTGGGCTCCTTCTGCATGATTTCTATCCAAATTCACAGACAGTTTGGTGAAGTTGCAATCC 1252
|||
Db 8981 CTGGGCTGTTACTTTCATGATGCCATCCAGATCCCCCGGCAGCTAGTGAAGTAGCAGCT 9040
|||
QY 1253 TTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTTCCAA 1297
|||
Db 9041 TTTGGAGGCAGTAATATTGAGCCTAGTGTTCGACAGCTGCTTTCCAA 9085
|||

RESULT 7

JS-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 290344.1
JS-09-976-594-93

Query Match 41.5%; Score 540.2; DB 4; Length 3915;
Best Local Similarity 64.3%; Pred. No. 2.3e-165;
Matches 827; Conservative 0; Mismatches 458; Indels 2; Gaps 1;

14	GTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAACATAAAGAACCT	73
14		
722	GTGCAACAGGAGAGGAGACACATCGGCGCTTTATGGAAGAAGTCAAGTCTCGGGGCCCC	781
74	GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGA	133
74		
782	TACATCTATTCTGTGCTGGAGTCAGCTCAGGCCCTTCTGTCCAGCACCCATTGAGGAG	841
134	CTAGAGAACTCTACAGGAGCCGACAGAGAGTGCCTCTGAGGAGAGAGAGCCAGAAATGTC	193
134		
842	TTAGAGGAGCCTCATTTCTGAGAGCAAGATACCTCCCGAAACAGCGGATCCAGAATCTC	901
194	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAGTGGGAAATTTGAACCTG	253
194		
902	AGCCGCTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACCTGTGGAGAAAGTTGACAGCC	961
254	CACCTCGCTGACTGGCAGAGAAATAAGATGAGACCCCTTGAAGAATCCAGGAACCTTCAA	313
254		
962	CGCTGTGTGGACACAGCACCGTCAATTTGAGCGGACTCTGGAGCAGCTCTTTGGAGATTCA	1020
314	GAGGCCAGGATGAGCTGGACCTCAAGCTGCGCAAGTGGAGGTGATCAAGGATCCTGG	373
1021	-GGGGCATGGAGGAACCTAAGCACTACTCTGAGCCAAGTGGAGGAGTCCGAGCCACTTGG	1079
374	CAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA	433
1080	GAGCCCAATTGGGGATCTCTTCAATTGATTCACTCCGACAGCACATCCAGGCTATTAAAGCTG	1139
434	CTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGCTCGC	493
1140	TTCAAAGAGAAATTCCTCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCAC	1199
494	CAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGAAGACCTG	553
1200	CAACTTGCCATTTCTGATGTGCATTTGTCAATGGAGAAATCCAGGCCCTTGAACAGATC	1259
554	AACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGGAGTGCATGAA	613
1260	AACGTCGATGGAACAACACTACAGGCGTCAGTTGATGAGAGGCTTAAGCAGTCCAGGAT	1319
614	GCCACAGGAGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCC	673
1320	GCCACCGGAGACTTTGGGCTGGGTACAGCACTTTCTCTCTCTCTCTCTCTCTCTCTCTCT	1379

RESULT 8

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313

QY	674	TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA	733
Db	1380	TGGGAAAGAGCAATTTACCCCAATAAAGTTCCTACTACATCAACACGAGCTCAGACC	1439
QY	734	ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAAT	793
Db	1440	ACATGCTGGGACCATCCCAAGATGACAGAGTTATACCAAAACCTAGCTGATCTGAACAAC	1499
QY	794	GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAGGCCCTTTGC	853
Db	1500	ATTAAGTTCTCAGCTTATCGCACATGCCATGAAACTCCGAGAGTCCAGAAAGCCCTGCGC	1559
QY	854	TTGGATCTCTTGAGCCTGTGAGCTGATGATGCCCTTGACCAGCACAACCTCAAGCAA	913
Db	1560	TTGGACCTGGTAACTTTAACCAACAGCCCTGGAATCTTCAATGAGCATGATCTGAGGCC	1619
QY	914	AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACACTATTATTATGACCGC	973
Db	1620	AGTGAGCACGTCATGGATGTTGGTAGAGGTCAATTCACCTGCCTGACTGCCTTATATGACGT	1679
QY	974	CTGGAGCAAGAGCACAACAATTTTGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAAC	1033
Db	1680	TTGGAGGAGGAAAGAGGCATCCTGTGTCACGTCGCCACTCTGTGGACATGAGCCCTCAAT	1739
QY	1034	TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTAAA	1093
Db	1740	TGGCTCCTCAATGTTTATGATAGTGTGCGAGCGGAAAGATGCGGGCATTTGCTTTAAG	1799
QY	1094	ACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG	1153
Db	1800	ACTGGCATTTGCATGCTTGTGTGGCACGGAAGTGAAGGAAAACTTCAGTACCTCTTCAGC	1859
QY	1154	CAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGGAGGCTGGCCCTCTTCTGCAATGAT	1213
Db	1860	CAAGTGGCAACTCAGGAGCAGCCAGTGTACCCAGCCACCTTGGTGTCTGCTTCATGAG	1919
QY	1214	TCTATCCAAATTCGAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGAGTAACATGAG	1273
Db	1920	GCCATTGAGGTGCCCCGTGAGTGGGTGAAGTGGCAGCCTTTGGGGGAGCAATGTGGAG	1979
QY	1274	CCAAGTGTCCGGAGCTGCTTCCAAATTT	1300
Db	1980	CCCAGTGTCCGTAGTTGCTTCCGTTTT	2006


```

; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
;
US-08-232-463-14

Query Match      3.8%; Score 50; DB 1; Length 7218;
Best Local Similarity 4.2%; Pred. No. 2.7e-05;
Matches 14; Conservative 191; Mismatches 131; Indels 0; Gaps 0;

QY 18 AGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAA 77
Db 1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322

QY 78 TCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAG 137
Db 1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262

QY 138 AGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTC 197
Db 1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202

QY 198 GCCTTCTACGAAAGCGGTGAGGAGGTCAATACTAGTGGGAAAATTTGAACCTGCACT 257
Db 1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142

QY 258 CCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTTCAAGAG 317
Db 1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082

QY 318 CCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTG 353
Db 1081 RRRRRRRRRRRRRRRATCGAAGCTCCCTCGACCTG 1046

RESULT 9
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match      3.4%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 4.2%; Pred. No. 0.0023;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

```

```

Best Local Similarity 54.7%; Pred. No. 0.0011;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 674 TGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 733
Db 1733 TGGGAGGAAGCTTACACAGCAGATGGATCAAGTACTTTCATCAACACGACGACACC 1792

QY 734 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 793
Db 1793 ACGTCTGGATCCACCCCGTGATGAGCGCCCTGAACCTGTCTCTGTGCAGAGGAGAGTGAA 1852

QY 794 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAG 834
Db 1853 GAGGACTGTCCAGAGAGACTAACAGACCCGAAAAAGCTGATG 1893

RESULT 10
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match      3.3%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0023;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 301 CCAGGAACCTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGTGAGGTGAT 360
Db 55 CAAGGAGTTGCACAGGTGGCGCACGACCTGGACGAGAGCTGGCATGGGTTTCAGGAGCG 114

QY 361 CAAGGGA--TCCTGGAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCT 417
Db 115 GCTGCCACTGGCCATGCAGACAGAGCGGCAACGGTTTGCAGGCGGTTCAGCAGCACAT 174

QY 418 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGT 477
Db 175 CAAAAGAACCAAGGCGCTTCGCGGGGAGATCCAGGCGCATGGGCGCGCTGGAGGAGGT 234

```


Db 971 TAGGAGGTTATGGACCAAGGACAAAGGTCAGGAGGATATGGACCAAGGACAAAGGTC 1030
QY 311 CAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCC 370
Db 1031 CAGGAGGATATGGACAGGTAGTGCACTCTGCAGCAGCAGCCGCGAGCAGGACCTGGACAAC 1090
QY 371 TGGCAGCCCGTGG 383
Db 1091 AAGGACCAGGAGG 1103

RESULT 13

US-08-317-844B-3
; Sequence 3, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1995 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: p6B
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1785
US-08-317-844B-3

Query Match 3.0%; Score 38.6; DB 2; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.061;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 11 GCAGTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACATAAGAA 70
Db 731 GCAGTGCAGCTGCAGCAGCCGCGCAGCAGGACCTGGACAACAAGGACCCCGAGGATATGGAC 790
QY 71 CCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAA 130
Db 791 CAGGACAACAAGGACCACCATCTGGACCCCGGTAGTGGCCGCTGCAGCAGCAGCCGCGCAGCAG 850

QY 131 GGACTAGAGAAACTCTACCAAGGAGCCCAAGAGAGTGCTCCTCTAGAGGAGAGAGCCAGAAAT 190
Db 851 GACCTGGAGGATATGGCCCTGGACAACAAGGACCCGAGGATATGGACCAAGCAAG 910
QY 191 GTCACTGGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAAC 250
Db 911 GACCATCTGGAGCAGGAGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 970
QY 251 CTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTT 310
Db 971 TAGGAGGTTATGGACCAAGGACAAAGGTCAGGAGGATATGGACCAAGGACAAAGGTC 1030
QY 311 CAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCC 370
Db 1031 CAGGAGGATATGGACCAAGGTCAGTGTGCACTCTGCAGCAGCAGCCGCGAGGACCTGGACAAC 1090
QY 371 TGGCAGCCCGTGG 383
Db 1091 AAGGACCAGGAGG 1103

RESULT 14

US-09-220-132-24
; Sequence 24, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 7672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-24

Query Match 3.0%; Score 38.4; DB 4; Length 7672;
Best Local Similarity 47.2%; Pred. No. 0.17;
Matches 117; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 375 AGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCAC 434
Db 6431 ACCCCGTGCGCTGCAACTCCTTGGAGAAATCAAAGCTTTGCGGAGGCCACGACGCCT 6490
QY 435 TTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTGCTCGCC 494
Db 6491 TCCGCTCCTCCTCAGCTCTGCGCAGGCTGACTTCAACACAGCTGGCCGAGTGGACCCGCC 6550
QY 495 AGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGA 554
Db 6551 AGATCAAGAGCTTCCGCGTAGCCTCCAAACCCCTACACCTGGTTTACCATGGAGGCCCTGG 6610
QY 555 ACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAG 614
Db 6611 AGGAGACCTTGAGGAACCTACAGAAAAATCATCAAGGAGAGGGAGCTGGAGCTGCAGAAGG 6670
QY 615 CCCACAGG 622
Db 6671 AACAGCGG 6678

RESULT 15

US-09-668-313A-3
; Sequence 3, Application US/09668313A

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

MM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 04:53:30 ; Search time 388.214 seconds
(without alignments)
15454.047 Million cell updates/sec

Title: US-09-845-416-14_COPY_900_2500
Effect score: 1601
Sequence: 1 gagctatgcctacacagg.....taccagtcttttagtgacct 1601

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1601	100.0	3446	10	US-09-845-416-14 Sequence 14, Appl
2	1590	99.3	4414	10	US-09-845-416-32 Sequence 32, Appl
3	1420.8	88.7	5417	15	US-10-149-736-39 Sequence 39, Appl
4	1193	74.5	3510	10	US-09-845-416-12 Sequence 12, Appl
5	1193	74.5	4476	10	US-09-845-416-31 Sequence 31, Appl
6	1102	68.8	4182	10	US-09-845-416-2 Sequence 2, Appl
7	1102	68.8	5149	10	US-09-845-416-27 Sequence 27, Appl
8	1091	68.1	8689	15	US-10-149-736-42 Sequence 42, Appl
9	1091	68.1	11443	15	US-10-149-736-44 Sequence 44, Appl
10	1091	68.1	12057	15	US-10-149-736-47 Sequence 47, Appl
11	1091	68.1	13957	9	US-09-782-378A-22 Sequence 22, Appl
12	1091	68.1	13957	9	US-09-880-107-2284 Sequence 2284, Ap
13	1091	68.1	13957	15	US-10-149-736-1 Sequence 1, Appli
14	1091	68.1	14069	12	US-10-342-887-434 Sequence 434, App
15	1091	68.1	14082	12	US-10-342-887-981 Sequence 981, App

16	1091	68.1	14082	15	US-10-341-434-108 Sequence 108, App
17	1087.8	67.9	11058	10	US-09-845-416-1 Sequence 1, Appli
18	1085.8	67.8	1991	10	US-09-845-416-3 Sequence 3, Appli
19	863	53.9	3858	10	US-09-845-416-9 Sequence 9, Appli
20	863	53.9	4825	10	US-09-845-416-29 Sequence 29, Appli
21	863	53.9	4848	10	US-09-845-416-35 Sequence 35, Appli
22	863	53.9	5060	10	US-09-845-416-36 Sequence 36, Appli
23	854.2	53.4	13815	15	US-10-149-736-2 Sequence 2, Appli
24	854	53.3	5339	15	US-10-149-736-40 Sequence 40, Appli
25	851.2	53.2	3531	10	US-09-845-416-10 Sequence 10, Appli
26	851.2	53.2	4498	10	US-09-845-416-30 Sequence 30, Appli
27	766.4	47.9	3999	10	US-09-845-416-6 Sequence 6, Appli
28	766.4	47.9	4966	10	US-09-845-416-28 Sequence 28, Appli
29	766.4	47.9	4990	10	US-09-845-416-34 Sequence 34, Appli
30	765	47.8	1667	10	US-09-845-416-7 Sequence 7, Appli
31	732.2	45.7	5462	15	US-10-149-736-41 Sequence 41, Appli
32	502.6	31.4	1821	10	US-09-845-416-13 Sequence 13, Appli
33	502.6	31.4	2169	10	US-09-845-416-4 Sequence 4, Appli
34	501	31.3	1434	10	US-09-845-416-15 Sequence 15, Appli
35	450	28.1	1340	10	US-09-845-416-11 Sequence 11, Appli
36	349.4	21.8	10302	9	US-09-782-378A-23 Sequence 23, Appli
37	349.4	21.8	10302	15	US-10-149-736-3 Sequence 3, Appli
38	349.4	21.8	16531	14	US-10-101-510-667 Sequence 667, App
39	327	20.4	327	15	US-10-149-736-8 Sequence 8, Appli
40	326.2	20.4	11096	15	US-10-149-736-4 Sequence 4, Appli
41	324	20.2	324	15	US-10-149-736-33 Sequence 33, Appli
42	322	20.1	333	15	US-10-149-736-10 Sequence 10, Appli
43	321	20.0	333	15	US-10-149-736-9 Sequence 9, Appli
44	245.4	15.3	5106	12	US-10-220-120-157 Sequence 157, App
45	190.2	11.9	256	9	US-09-864-761-21956 Sequence 21956, A

ALIGNMENTS

RESULT 1
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-14

Query Match	100.0%;	Score 1601;	DB 10;	Length 3446;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1601;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GAGCTATGCCTACACACAGGCTGCTTATGTCACCCCTCTGACCCCTACACGGAGCCATT	60	
Db	900	GAGCTATGCCTACACACAGGCTGCTTATGTCACCCCTCTGACCCCTACACGGAGCCATT	959	
Qy	61	TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG	120	
Db	960	TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG	1019	
Qy	121	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC	180	
Db	1020	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC	1079	
Qy	181	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTATCATGTGTGAAGTGTGAAGA	240	

1080 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGSGCCGGTTGG 300
1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGSGCCGGTTGG 1199
301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAAGATGAAGA 360
1200 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAAGATGAAGA 1259
361 AACTGAAGTACAAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
1260 AACTGAAGTACAAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
421 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGATGGATCTCCAGAAATCGAAACTG 480
1320 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGATGGATCTCCAGAAATCGAAACTG 1379
481 AAAGAGTTGATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAAGAAATGGAGGAAGAG 540
1380 AAAGAGTTGATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGAG 1439
541 CCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACATAAAGTGCCTTCAA 600
1440 CCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACATAAAGTGCCTTCAA 1499
601 GAAAGATCTAGAAACAAGAACAAAGTCAAGTCAAGTCAATCTCTCACTCACATGGTGGTGGTAGTT 660
1500 GAAAGATCTAGAAACAAGAACAAAGTCAAGTCAAGTCAATCTCTCACTCACATGGTGGTGGTAGTT 1559
661 GATGAATCTAGTGGAGATCACGCCAATCGCTGCTTTGGAAAGAACAACTTAAGTATTGGGA 720
1560 GATGAATCTAGTGGAGATCACGCCAATCGCTGCTTTGGAAAGAACAACTTAAGTATTGGGA 1619
721 GATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTCTCTTTTACAAACATC 780
1620 GATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTCTCTTTTACAAACATC 1679
781 CTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTTCAGAA 840
1680 CTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTTCAGAA 1739
841 AAAGAAGATGCAATGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAATGTTA 900
1740 AAAGAAGATGCAATGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAATGTTA 1799
901 TCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCATG 960
1800 TCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCATG 1859
961 GGCAAACTGTATTCACTCAAAACAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACC 1020
1860 GGCAAACTGTATTCACTCAAAACAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACC 1919
1021 CAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAAA 1080
1920 CAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAAA 1979
1081 CTTGAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGAT 1140
1980 CTTGAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGAT 2039
1141 GAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCTCGGACGCCCGTGGGC 1200
2040 GAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCTCGGACGCCCGTGGGC 2099
1201 GATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAA 1260
2100 GATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGSCACTTCGAGGAGAA 2159
1261 ATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGTCTGCCAGCTTACCCT 1320

2150 ATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCATATGACCTTGTCTGCCAGCTTACCCT 2219
1321 TTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGG 1380
2220 TTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGG 2279
1381 AAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCCAACAGGGAC 1440
2280 AAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCCAACAGGGAC 2339
1441 TTTGGTCCAGCATCTCAGCACATTTCTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGCC 1500
2340 TTTGGTCCAGCATCTCAGCACATTTCTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGCC 2399
1501 ATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCACGAGACTCAAAACAACTTGTCTGGAC 1560
2400 ATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCACGAGACTCAAAACAACTTGTCTGGAC 2459
1561 CATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
2460 CATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2500

RESULT 2
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 99.3%; Score 1590; DB 10; Length 4414;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGTACCCCTACACGAGCCCCATT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGTACCCCTACACGAGCCCCATT 1716
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCCGGTTGG 300
Db 1897 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCCGGTTGG 1956
QY 301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

||||| 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTTAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
||||| 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATC-GAAACT 479
||||| 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 2136
||||| 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGGAAAAATGGAGGAAGA 539
||||| 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGGAAAAATGGAGGAAGA 2196
||||| 540 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATTAAGTGTCTTCA 599
||||| 2197 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATTAAGTGTCTTCA 2256
||||| 600 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACTGATGGTGGTGTAGT 659
||||| 2257 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACTGATGGTGGTGTAGT 2316
||||| 660 TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 719
||||| 2317 TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
||||| 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT 779
||||| 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT 2436
||||| 780 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGATGGCTTTCAGA 839
||||| 2437 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGATGGCTTTCAGA 2496
||||| 840 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
||||| 2497 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 2556
||||| 900 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
||||| 2557 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616
||||| 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAAGTCAGTGAC 1019
||||| 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAAGTCAGTGAC 2676
||||| 1020 CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
||||| 2677 CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2736
||||| 1080 ACTTGAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1139
||||| 2737 ACTTGAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 2796
||||| 1140 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGCGAGCCCGTGGG 1199
||||| 2797 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGCGAGCCCGTGGG 2856
||||| 1200 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 1259
||||| 2857 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 2916
||||| 1260 AATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1319
||||| 2917 AATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 2976
||||| 1320 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 1379
||||| 2977 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 3036
||||| 1380 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGTCATGAAGCCCAAGGGA 1439
||||| 3037 GAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGTCATGAAGCCCAAGGGA 3096
||||| 1440 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 1499
|||||

Db 3097 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGAGAGAGC 3156
QY 1500 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTCTGCTGGGA 1559
||||| 3157 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTCTGCTGGGA 3216
QY 1560 CCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 1601
||||| 3217 CCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 3258
Db
RESULT 3
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39
Query Match 88.7%; Score 1420.8; DB 15; Length 5417;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 2; Indels 157; Gaps 2;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGTATGGAGAG 120
||||| 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGTATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCCTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAAGTGAAGTGGAAACAGGAAAAATTCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAAGTGAAGTGGAAACAGGAAAAATTCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGCTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGCTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGATGACTGGCTTAACAAAAACAGAGAAAGAAACAGGAAAAATGGAGGAAGA 539

Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTTGAAGACCTPAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTPAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
Db 1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 780 CCTTCTCAAATGSCAACGCTCTTACTGAGAACAGTGCCTTTTATTAGTGCATGGCTTTCAGA 839
Db 1879 CCTTCTCAAATGSCAACGCTCTTACTGAGAACAGTGCCTTTTATTAGTGCATGGCTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 1998
QY 900 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
Db 2119 CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1080 ACTTGAAAAGAGTACAGCACAGA----- 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTTCAGAGGCTGTCAACCACTCAGCCATCACTAAC 2238
QY 1103 ----- 1102
Db 2239 ACAGACAACGTGTAATGGAAACAGTAACTACGGTGACCAAGGGAACAGATCCTGGTAAA 2298
QY 1103 -----C 1103
Db 2299 GCATGCTCAAGAGGAACTTCCACCACCTCCCCAAAAGAGAGCAGATTACTGTGGA 2358
QY 1104 CCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGGCCCA 1163
Db 2359 TCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGGCCCA 2418
QY 1164 AGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCA 1223
Db 2419 AGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCA 2478
QY 1224 AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTCGCGCTCTGAAAGAGAACGT 1283
Db 2479 AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTCGCGCTCTGAAAGAGAACGT 2538
QY 1284 GAGCACGTCGAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTA 1343
Db 2539 GAGCACGTCGAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTA 2598
QY 1344 TAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCTGAGGTGGCCGTCGA 1403
Db 2599 TAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCTGAGGTGGCCGTCGA 2658
QY 1404 GGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTT 1463

Db 2659 GGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTT 2718
QY 1464 TCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTA 1523
Db 2719 TCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTA 2778
QY 1524 CTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTA 1583
Db 2779 CTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTA 2838
QY 1584 CCAGTCTTTAGCTGACCT 1601
Db 2839 CCAGTCTTTAGCTGACCT 2856

RESULT 4
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 74.5%; Score 1193; DB 10; Length 3510;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

QY 1 GAGTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCCATT 60
Db 900 GAGTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCCATT 959
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTGCAAGCAAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
Db 1080 TGCTGAGGACACATTGCAAGCAAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1139
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGCCCGGGTTGG 300
Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGCCCGGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539

b 1380 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAATGGAGGAAGA 1439
y 540 GCCTCTTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACACATAAAGGTGCTTCA 599
b 1440 GCCTCTTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAAACACATAAAGGTGCTTCA 1499
y 600 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 659
b 1500 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 1559
y 660 TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 719
b 1560 TGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1619
y 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTGGTCTTTTACAAGACAT 779
b 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTGGTCTTTTACAAGACAG 1679
y 780 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAG 839
b 1680 TTCTGACCAGTGGAGCGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCT 1739
y 840 AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 899
b 1740 GAAAGATGATGAATTAACCGCGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAA 1799
y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAACCGGATCTAGAAAAGAAAAGCAATCCAT 959
b 1800 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCAT 1859
y 960 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACACTGAAG 1005
b 1860 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1919
y 1006 AATAAGTCAGTACCCAGAACGGAAGCATGGCTGGA----- 1043
b 1920 ACTCTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCT 1979
y 1044 -----TAACTTTGCCCGGTGGGATAATTTAGTCCAAAACCTTGAA----- 1086
b 1980 TCTACGAAAGCAGGCTGAGGAGGTCAAATCTGAGTGGGAAAAATTGAACCTGCACCTCCGC 2039
y 1087 -----AAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAATCTCAAGAGGCCAC 1136
b 2040 TGAAGTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAATCTCAAGAGGCCAC 2099
y 1137 GGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1196
b 2100 GGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2159
y 1197 GGGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGG 1256
b 2160 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGG 2219
y 1257 AGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGGCTCGCCAGCTTAC 1316
b 2220 AGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGGCTCGCCAGCTTAC 2279
y 1317 CACTTTGGGCATTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAG 1376
b 2280 CACTTTGGGCATTCAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAG 2339
y 1377 ATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAG 1436
b 2340 ATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAG 2399
y 1437 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1496
b 2400 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 2459
y 1497 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAAACTTGCTG 1556

Db 2460 AGCCATCTGCGCCAAACAAAGTGCCTTACTATATCAACCAGAGACTCAAAACAACCTTGCTG 2519
QY 1557 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
Db 2520 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2564
RESULT 5
US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31
Query Match 74.5%; Score 1193; DB 10; Length 4476;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GAGCTATGCGCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACCGAGCCCAT 60
Db 1656 GAGCTATGCGCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACCGAGCCCAT 1715
QY 61 TCCTTCACAGCATTTGGAAGTCTCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG 120
Db 1716 TCCTTCACAGCATTTGGAAGTCTCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG 1775
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
Db 1776 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1835
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1836 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1895
QY 241 CCAGTTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1896 CCAGTTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1955
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1956 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2015
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2016 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2075
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 2076 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2135
QY 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
Db 2136 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 2195
QY 540 GCCTCTTTGGACCTGATCTTGAAGACCTAAAAACGCAAGTACAAACAACATAAGGTGCTTCA 599
Db 2196 GCCTCTTTGGACCTGATCTTGAAGACCTAAAAACGCAAGTACAAACAACATAAGGTGCTTCA 2255
QY 600 AGAAGATCTAGAAACAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659

db 2256 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 2315
2y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGG 719
db 2316 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGG 2375
2y 720 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
db 2376 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 2435
2y 780 CCTTCTCAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 839
db 2436 TTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 2495
2y 840 AAAAGAAGATGAGTGAACAAGATTACACAACACTGGCTTTAAAGATCAAAATGAAATGTT 899
db 2496 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAA 2555
2y 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 959
db 2556 GCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTGAAAACCTAAAGAACCTTAAATCAT 2615
2y 960 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG 1005
db 2616 GAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGCCTTTGGAAGGACTAGAGAA 2675
2y 1006 AATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGA----- 1043
db 2676 ACTTACGAGGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 2735
2y 1044 -----TAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAAACTTGAA----- 1086
db 2736 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAAACCTGCACTCCGC 2795
2y 1087 -----AAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1136
db 2796 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCAC 2855
2y 1137 GGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1196
db 2856 GGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2915
2y 1197 GGGCGATCTCCTCAATGACTCTCTCCAAGATCACCTCGAAGAGTCAAGGCACCTTCGAGG 1256
db 2916 GGGCGATCTCCTCAATGACTCTCTCCAAGATCACCTCGAAGAGTCAAGGCACCTTCGAGG 2975
2y 1257 AGAAATTGCGCCTCTGAAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTAC 1316
db 2976 AGAAATTGCGCCTCTGAAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTAC 3035
2y 1317 CACTTTGGGCATTCAAGCTCTCACCGGTAAACCTCAGCACCTCTGGAAGACCTGAAACACCAG 1376
db 3036 CACTTTGGGCATTCAAGCTCTCACCGGTAAACCTCAGCACCTCTGGAAGACCTGAAACACCAG 3095
2y 1377 ATGGAAGCTTTCGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 1436
db 3096 ATGGAAGCTTTCGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 3155
2y 1437 GGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1496
db 3156 GGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 3215
2y 1497 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTGCTG 1556
db 3216 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTGCTG 3275
2y 1557 GGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCT 1601
db 3276 GGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCT 3320

US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 68.8%; Score 1102; DB 10; Length 4182;
Best Local Similarity 96.9%; Pred. No. 3.9e-312;
Matches 1134; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCATT 60
Db 900 GAGCTATGCCTACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCATT 959

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019

QY 121 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1079

QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139

QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1199

QY 301 TAATATTCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAAAATTCAGAAGATGAAGA 1259

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAAATTCGAAGATGGGAATGCCTCAGGGTAGC 1319

QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379

QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGAAAAATGGAGGAAGA 539
Db 1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAAAGAAAAATGGAGGAAGA 1439

QY 540 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1440 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1499

QY 600 AGRAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTCACTCAATGGTGGTGGTAGT 659
Db 1500 AGRAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTCACTCAATGGTGGTGGTAGT 1559

QY 660 TGATGAATCTAGTGGAGATCAGCCTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGG 719
Db 1560 TGATGAATCTAGTGGAGATCAGCCTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGG 1619

QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779

|||||
1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGTGGTTCCTTTACAAGACAT 1679
|||
780 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 839
|||||
1680 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTCAGA 1739
|||||
840 AAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
|||||
1740 AAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1799
|||||
900 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|||||
1800 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1859
|||||
960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATPAAAGTCAGTGAC 1019
|||||
1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAAATPAAAGTCAGTGAC 1919
|||||
1020 CCAGAAGACGGAGCAGTGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
|||||
1920 CCAGAAGACGGAGCAGTGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1979
|||||
1080 ACTTGAAGAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1139
|||||
1980 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2039
|||||
1140 TGAGCTGGACCTCAAGCTGGGCCAAGCTGA 1169
|||
2040 AAAGTTTCTTGCTTGCTTACAGAAGCTGA 2069
|||||

RESULT 7
IS-09-845-416-27
Sequence 27, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 5149
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-845-416-27

Query Match 68.8%; Score 1102; DB 10; Length 5149;
Best Local Similarity 96.9%; Pred. No. 4.5e-312;
Matches 1134; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
b 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 1716
Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTTGATGGAGAG 120
b 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTTGATGGAGAG 1776
Y 121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTC 180
b 1777 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTC 1836
Y 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
b 1837 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1896
|||||

QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
|||
Db 1897 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1956
|||
QY 301 TAATATTCTACAAATTGGGAAGTAAAGTCTGATTGGAAACAGSAAAAATTATCAGAAGATGAAGA 360
|||
Db 1957 TAATATTCTACAAATTGGGAAGTAAAGTCTGATTGGAAACAGSAAAAATTATCAGAAGATGAAGA 2016
|||
QY 361 AACTGAAGTACAAAGACAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 2017 AACTGAAGTACAAAGACAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
|||
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
|||
Db 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2136
|||
QY 480 GAAAGACTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA 539
|||
Db 2137 GAAAGACTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA 2196
|||
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAAGTGTCTTCA 599
|||
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAAGTGTCTTCA 2256
|||
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 659
|||
Db 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 2316
|||
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 719
|||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 2376
|||
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
|||
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 2436
|||
QY 780 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTTCAGA 839
|||
Db 2437 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTTCAGA 2496
|||
QY 840 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 899
|||
Db 2497 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 2556
|||
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|||
Db 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616
|||
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCACACACTGAAGAAATPAAAGTCAGTGAC 1019
|||
Db 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCACACACTGAAGAAATPAAAGTCAGTGAC 2676
|||
QY 1020 CCAGAAGACGGAGCAGTGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
|||
Db 2677 CCAGAAGACGGAGCAGTGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2736
|||
QY 1080 ACTTGAAGAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1139
|||
Db 2737 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCCCTTGGACCTGGA 2796
|||
QY 1140 TGAGCTGGACCTCAAGCTGGGCCAAGCTGA 1169
|||
Db 2797 AAAGTTTCTTGCTTGCTTACAGAAGCTGA 2826
|||

RESULT 8
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

```

; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/311126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; TYPE: DNA
; LENGTH: 8689
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match      68.1%; Score 1091; DB 15; Length 8689;
Best Local Similarity 99.9%; Pred. No. 1.1e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
Db      1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1158

QY      61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTCGSCAGTTTCATTGATGGAGAG 120
Db      61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTCGSCAGTTTCATTGATGGAGAG 1218

QY      121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTTTTC 180
Db      121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTTTTC 1278

QY      1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTTTTC 1278
Db      1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTTTTC 1278

QY      181 TGCTGAGGACACATTCGAAGCACACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db      181 TGCTGAGGACACATTCGAAGCACACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY      1279 TGCTGAGGACACATTCGAAGCACACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
Db      1279 TGCTGAGGACACATTCGAAGCACACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY      241 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db      241 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398

QY      1339 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398
Db      1339 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398

QY      301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA 360
Db      301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA 1458

QY      1399 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA 1458
Db      1399 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA 1458

QY      361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db      361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518

QY      1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
Db      1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518

QY      421 TAGCATGGAACCAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC-GAAACT 479
Db      421 TAGCATGGAACCAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578

QY      1519 TAGCATGGAACCAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
Db      1519 TAGCATGGAACCAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578

QY      480 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAATGGAGGAAGA 539
Db      480 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAATGGAGGAAGA 1638

QY      1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAATGGAGGAAGA 1638
Db      1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGAAAGAACAGGAAAATGGAGGAAGA 1638

QY      540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 599
Db      540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 1698

QY      1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 1698
Db      1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 1698

QY      600 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 659
Db      600 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1758

QY      1699 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1758
Db      1699 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1758

QY      660 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db      660 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

QY      1759 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
Db      1759 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

QY      720 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db      720 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

QY      1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878
Db      1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878
```

```

QY      780 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATTAGTCATGGCTTTTTCAGA 839
Db      1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATTAGTCATGGCTTTTTCAGA 1938

QY      840 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 899
Db      1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 1998

QY      900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db      1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058

QY      960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db      2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

QY      1020 CCAGAAGACGGAAGCATGGCTGGATAACCTTTGCCCGGCTGTTGGGATAATTTAGTCCAAAA 1079
Db      2119 CCAGAAGACGGAAGCATGGCTGGATAACCTTTGCCCGGCTGTTGGGATAATTTAGTCCAAAA 2178

QY      1080 ACTTGAAGAGAGTACAGCACAGA 1102
Db      2179 ACTTGAAGAGAGTACAGCACAGA 2201

RESULT 9
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/311126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-44

Query Match      68.1%; Score 1091; DB 15; Length 11443;
Best Local Similarity 99.9%; Pred. No. 1.3e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 60
Db      1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT 1158

QY      61 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db      1159 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1218

QY      121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db      1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

QY      181 TGCTGAGGACACATTTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db      1279 TGCTGAGGACACATTTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY      241 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db      1339 CCAGTTTCATACTATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398
```



```

>Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
>b 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
>Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
>b 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 1518
>Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 479
>b 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
>Y 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539
>b 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1638
>Y 540 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAAGGTGCTTCA 599
>b 1639 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAAGGTGCTTCA 1698
>Y 600 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 659
>b 1699 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 1758
>Y 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
>b 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
>Y 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGTGGTTCTTTTACAAAGACAT 779
>b 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGTGGTTCTTTTACAAAGACAT 1878
>Y 780 CCITTCTCAAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTTAGTGGCATGGCTTTCAGA 839
>b 1879 CCITTCTCAAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTTAGTGGCATGGCTTTCAGA 1938
>Y 840 AAAAGAAGATGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
>b 1939 AAAAGAAGATGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
>Y 900 ATCAAGTCTTCAAAAACTGSCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
>b 1999 ATCAAGTCTTCAAAAACTGSCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
>Y 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1019
>b 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 2118
>Y 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAATTTAGTCCAAAA 1079
>b 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTGGGATAATTTAGTCCAAAA 2178
>Y 1080 ACTTGAAAAAGAGTACAGCACAGA 1102
>b 2179 ACTTGAAAAAGAGTACAGCACAGA 2201
```

ESULT 10
IS-10-149-736-47
Sequence 47, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06

```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 12057
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match      68.1%; Score 1091; DB 15; Length 12057;
Best Local Similarity 99.9%; Pred. No. 1.3e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCATTT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCATTT 1158
QY 61 TCCTTCACACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGAAGA 1338
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATGTAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAAACT 479
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAACAACTAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 659
Db 1699 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGTGGGTTCTTTTACAAAGACAT 779
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGTGGGTTCTTTTACAAAGACAT 1878
QY 780 CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 839
Db 1879 CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1938
QY 840 AAAAGAAGATGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAAGATTTCACACAACCTGGCTTTTAAAGATCAAAAATGAAATGTT 1998

```

QY 900 ATCAAGTCTTCAAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 2178
QY 1080 ACTTGAAAAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 11
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782.378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match 68.1%; Score 1091; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGCGAGTTTCATTTGAGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGCGAGTTTCATTTGAGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518

QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATC-GAAACT 479
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 480 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAACAAAGGAAAATGGAGGAAGA 539
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAACAAAGGAAAATGGAGGAAGA 1638
QY 540 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
Db 1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 659
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1758
QY 660 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTTTGGAAGAACAACTTAAGGTATTGGG 719
Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTTTGGAAGAACAACTTAAGGTATTGGG 1818
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 779
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1878
QY 780 CCTTCTCAAATGGCAACGTCTTACTGAAACAGTGCCTTTTAAAGATCAAAATGAAATGTT 839
Db 1879 CCTTCTCAAATGGCAACGTCTTACTGAAACAGTGCCTTTTAAAGATCAAAATGAAATGTT 1938
QY 840 AAAAGAAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
Db 1939 AAAAGAAGATGCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 900 ATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 1999 ATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 960 GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 1019
Db 2059 GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 2178
QY 1080 ACTTGAAAAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 12
US-09-880-107-2284
; Sequence 2284, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2284
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 M18533
JS-09-880-107-2284

Query Match 68.1%; Score 1091; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACTCTGACCCTACACGGAGCCCAT 60
1099 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACTCTGACCCTACACGGAGCCCAT 1158
61 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
1159 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1278
181 TGCTGAGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 300
1339 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 1398
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGAGAAAATTTATCAGAAGATGAAGA 360
1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGAGAAAATTTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGAGATGGGAATGCCCTCAGGGTAGC 420
1459 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGAGATGGGAATGCCCTCAGGGTAGC 1518
421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
480 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAACAGGAAATGGAGGAAGA 539
1579 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAACAGGAAATGGAGGAAGA 1638
540 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCAAGTACAACACATAAGGTGCTTCA 599
1639 GCCTCTTGGAACCTGATCTTTGAAGACCTTAAACCGCAAGTACAACACATAAGGTGCTTCA 1698
600 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACTCACTCACTCACTCACT 659
1699 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACTCACTCACTCACTCACT 1758
660 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
1759 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
720 AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
1819 AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
780 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTAAAGATCAAAATGAATGTT 839
1879 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTAAAGATCAAAATGAATGTT 1938
840 AAAAGAAGATGCAGTGAACAAAGATTCACACAACTGGCTTTAAAGATCAAAATGAATGTT 899
1939 AAAAGAAGATGCAGTGAACAAAGATTCACACAACTGGCTTTAAAGATCAAAATGAATGTT 1998
900 ATCAAGTCTTCAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
1999 ATCAAGTCTTCAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
960 GGGCAAACTGTATTCACCTCAAAACAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1019

Db 2059 GGGCAAACTGTATTCACCTCAAAACAGATCTTCTTTCACACACTGAAGATAAGTCAGTGAC 2118
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1079
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 2178
QY 1080 ACTTGAAGAGAGTACAGCACAGA 1102
Db 2179 ACTTGAAGAGAGTACAGCACAGA 2201

RESULT 13
US-10-149-736-1
; Sequence 1, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-149-736-1

Query Match 68.1%; Score 1091; DB 15; Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
1 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACTCTGACCCTACACGGAGCCCAT 60
1099 GAGCTATGCCCTACACACAGGCTGCTTATGTACCACTCTGACCCTACACGGAGCCCAT 1158
61 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
1159 TCCTTCACAGCATTTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1278
181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 300
1339 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 1398
301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGAGAAAATTTATCAGAAGATGAAGA 360
1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGAGAAAATTTATCAGAAGATGAAGA 1458
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
480 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAACAGGAAATGGAGGAAGA 539
1579 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAACAGGAAATGGAGGAAGA 1638

QY 540 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
|
|
|
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698
|
|
|
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
|
|
|
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758
|
|
|
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGG 719
|
|
|
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGG 1818
|
|
|
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 779
|
|
|
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
|
|
|
QY 780 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
|
|
|
Db 1879 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938
|
|
|
QY 840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
|
|
|
Db 1939 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
|
|
|
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|
|
|
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
|
|
|
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 1019
|
|
|
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 2118
|
|
|
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
|
|
|
Db 2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 14

US-10-342-887-434
; Sequence 434, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 434
; LENGTH: 14069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-434

Query Match

68.1%; Score 1091; DB 12; Length 14069;

Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCTACACGGAGCCCAT 60
|
|
|
Db 1211 GAGCTATGCCTACACACAGGCTGCTTATGTCTACACCTCTGACCTACACGGAGCCCAT 1270
|
|
|
QY 61 TCCTTCACAGCATTTGGAAAGCTCCTGAAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG 120
|
|
|
Db 1271 TCCTTCACAGCATTTGGAAAGCTCCTGAAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG 1330
|
|
|
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATTCGTGGCTTCTTTC 180
|
|
|
Db 1331 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATTCGTGGCTTCTTTC 1390
|
|
|
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGTGTAAGA 240
|
|
|
Db 1391 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGTGTAAGA 1450
|
|
|
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
|
|
|
Db 1451 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1510
|
|
|
QY 301 TAATATTTCAAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAAAATATCAGAAGATGAAGA 360
|
|
|
Db 1511 TAATATTTCAAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAAAATATCAGAAGATGAAGA 1570
|
|
|
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
|
|
|
Db 1571 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1630
|
|
|
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 479
|
|
|
Db 1631 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1690
|
|
|
QY 480 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 539
|
|
|
Db 1691 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1750
|
|
|
QY 540 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 599
|
|
|
Db 1751 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1810
|
|
|
QY 600 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 659
|
|
|
Db 1811 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 1870
|
|
|
QY 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAGAACAACTTAAGTATTGGG 719
|
|
|
Db 1871 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAGAACAACTTAAGTATTGGG 1930
|
|
|
QY 720 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 779
|
|
|
Db 1931 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1990
|
|
|
QY 780 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 839
|
|
|
Db 1991 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 2050
|
|
|
QY 840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 899
|
|
|
Db 2051 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 2110
|
|
|
QY 900 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
|
|
|
Db 2111 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2170
|
|
|
QY 960 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 1019
|
|
|
Db 2171 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAGTGAC 2230
|
|
|
QY 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
|
|
|

Db 2231 CCAGAAAGCGGAAGCATGGCTGGATAACTTTGGCCCGGTGTTGGGATAAATTAGTCCAAAA 2290
Y 1080 ACTTGAAAGAGGTACAGCACAGA 1102
Db 2291 ACTTGAAAGAGGTACAGCACAGA 2313
RESULT 15
JS-10-342-887-981
; Sequence 981, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 981
; LENGTH: 14082
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-10-342-887-981
Query Match 68.1%; Score 1091; DB 12; Length 14082;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
Db 1224 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1283
Y 61 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1284 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1343
Y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCTGGCTTCTTTC 180
Db 1344 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCTGGCTTCTTTC 1403
Y 181 TGCTGAGGACACATTGCAAGGACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1404 TGCTGAGGACACATTGCAAGGACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1463
Y 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGCCCGGTTGG 300
Db 1464 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGCCCGGTTGG 1523
Y 301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1524 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1583
Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 420
Db 1584 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCTCAGGGTAGC 1643
Y 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAATC-GAAACT 479
Db 1644 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAATCAGAAACT 1703

Qy 480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGAAAAATGGAGGAAGA 539
Db 1704 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGAAAAATGGAGGAAGA 1763
Qy 540 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA 599
Db 1764 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA 1823
Qy 600 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
Db 1824 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1883
Qy 660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAAGGTATGGG 719
Db 1884 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAAGGTATGGG 1943
Qy 720 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db 1944 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACAT 2003
Qy 780 CCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 839
Db 2004 CCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCAGA 2063
Qy 840 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 899
Db 2064 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 2123
Qy 900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
Db 2124 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2183
Qy 950 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db 2184 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2243
Qy 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1079
Db 2244 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2303
Qy 1080 ACTTGAAAAGAGTACAGCACAGA 1102
Db 2304 ACTTGAAAAGAGTACAGCACAGA 2326

Search completed: April 5, 2004, 16:34:31
Job time : 393.214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 04:53:30 ; Search time 315.469 Seconds
(without alignments)
15454.047 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100

Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgttccaatttg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1301	100.0	1821	10	US-09-845-416-13
2	1301	100.0	2169	10	US-09-845-416-4
3	1301	100.0	3510	10	US-09-845-416-12
4	1301	100.0	3531	10	US-09-845-416-10
5	1301	100.0	3858	10	US-09-845-416-9
6	1301	100.0	3999	10	US-09-845-416-6
7	1301	100.0	4182	10	US-09-845-416-2
8	1301	100.0	4476	10	US-09-845-416-31
9	1301	100.0	4498	10	US-09-845-416-30
10	1301	100.0	4825	10	US-09-845-416-29
11	1301	100.0	4848	10	US-09-845-416-35
12	1301	100.0	4966	10	US-09-845-416-28
13	1301	100.0	4990	10	US-09-845-416-34
14	1301	100.0	5060	10	US-09-845-416-36
15	1301	100.0	5149	10	US-09-845-416-27

16	1301	100.0	5339	15	US-10-149-736-40	Sequence 40, Appl
17	1301	100.0	5462	15	US-10-149-736-41	Sequence 41, Appl
18	1301	100.0	8689	15	US-10-149-736-42	Sequence 42, Appl
19	1301	100.0	11058	10	US-09-845-416-1	Sequence 1, Appl
20	1301	100.0	11443	15	US-10-149-736-44	Sequence 44, Appl
21	1301	100.0	12057	15	US-10-149-736-47	Sequence 47, Appl
22	1301	100.0	13957	9	US-09-782-378A-22	Sequence 22, Appl
23	1301	100.0	13957	9	US-09-880-107-2284	Sequence 2284, Ap
24	1301	100.0	13957	15	US-10-149-736-1	Sequence 1, Appl
25	1301	100.0	14069	12	US-10-342-887-434	Sequence 434, App
26	1301	100.0	14082	12	US-10-342-887-981	Sequence 981, App
27	1301	100.0	14082	15	US-10-341-434-108	Sequence 108, App
28	1135.2	87.3	13815	15	US-10-149-736-2	Sequence 2, Appl
29	1017.6	78.2	3446	10	US-09-845-416-14	Sequence 14, Appl
30	1017.6	78.2	4414	10	US-09-845-416-32	Sequence 32, Appl
31	1016.6	78.1	5417	15	US-10-149-736-39	Sequence 39, Appl
32	1016	78.1	1434	10	US-09-845-416-15	Sequence 15, Appl
33	567.2	43.6	11096	15	US-10-149-736-4	Sequence 4, Appl
34	553.8	42.6	10302	9	US-09-782-378A-23	Sequence 23, Appl
35	553.8	42.6	10302	15	US-10-149-736-3	Sequence 3, Appl
36	551	42.4	5106	12	US-10-220-120-157	Sequence 157, App
37	545	41.9	16531	14	US-10-101-510-667	Sequence 667, App
38	472	36.3	887	15	US-10-149-736-35	Sequence 35, Appl
39	324	24.9	324	15	US-10-149-736-33	Sequence 33, Appl
40	289	22.2	387	15	US-10-149-736-32	Sequence 32, Appl
41	216	16.6	216	15	US-10-149-736-34	Sequence 34, Appl
42	114	8.8	114	15	US-10-149-736-45	Sequence 45, Appl
C 43	102.8	7.9	275	14	US-10-029-386-16813	Sequence 16813, A
C 44	102.8	7.9	587	14	US-10-029-386-3113	Sequence 3113, Ap
45	54.2	4.2	449	10	US-09-918-995-24084	Sequence 24084, A

ALIGNMENTS

RESULT 1

US-09-845-416-13
; Sequence 13, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-13

Query Match	100.0%	Score 1301;	DB 10;	Length 1821;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1301;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGACTTTCCAGCAGTTCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATTGAA	60	
Db	103	CGACTTTCCAGCAGTTCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATTGAA	162	
QY	61	AACTAAGAACCTGTATCATGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA	120	
Db	163	AACTAAGAACCTGTATCATGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA	222	
QY	121	GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCCGAGAGCTGCCTCCTGAGGAGAG	180	
Db	223	GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCCGAGAGCTGCCTCCTGAGGAGAG	282	
QY	181	AGCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGCTCAATACTAGTGGA	240	

Db: 283 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGA 342
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 343 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 402
QY 301 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 403 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 462
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 463 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 522
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 523 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 582
QY 481 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 583 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 642
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCTCGAGGACCGAGTCAG 600
Db 643 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCTCGAGGACCGAGTCAG 702
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 703 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 762
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 763 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 822
QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 823 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 882
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCA 840
Db 883 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAGACTGCA 942
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGACCCAGCA 900
Db 943 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGACCCAGCA 1002
QY 901 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAATTTAATTTGTTGACCCAC 960
Db 1003 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAATTTAATTTGTTGACCCAC 1062
QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1020
Db 1063 TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1122
QY 1021 TATGTGCTGAAGTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db 1123 TATGTGCTGAAGTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1182
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG 1140
Db 1183 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG 1242
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCCT 1200
Db 1243 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCCT 1302
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCAATCCTTTGGGGG 1260
Db 1303 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCAATCCTTTGGGGG 1362
QY 1261 CAGTAACATTGAGCAAGTGTCCGGAGCTGCTTCCAATTTG 1301

Db: 1363 CAGTAACATTGAGCAAGTGTCCGGAGCTGCTTCCAATTTG 1403
RESULT 2
US-09-845-416-4
; Sequence 4, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-4
Query Match 100.0%; Score 1301; DB 10; Length 2169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 451 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 510
QY 61 AACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 511 AACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAATATTTCTGACAGAGCA 570
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 571 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 630
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 240
Db 631 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 690
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 691 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 750
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 751 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 810
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 811 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 870
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 871 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 930
QY 481 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 931 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 990
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCGCTCGAGGACCGAGTCAG 600
Db 991 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCGCTCGAGGACCGAGTCAG 1050
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 1051 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

1111 TGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCTACTATATCAACCA 1170
721 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCACTCTTTAGC 780
1171 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCACTCTTTAGC 1230
781 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
1231 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1290
841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCTCAGCTGTCATGTGATGCCCTTGACCCAGCA 900
1291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCTCAGCTGTCATGTGATGCCCTTGACCCAGCA 1350
901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGAGTATTAATTTGTTGACCCAC 960
1351 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGAGTATTAATTTGTTGACCCAC 1410
961 TATTTATGACCGCCCTGGAGCAAGACCAACAATTTTGGTCAACGTCCTCTCTGCGTGA 1020
1411 TATTTATGACCGCCCTGGAGCAAGACCAACAATTTTGGTCAACGTCCTCTCTGCGTGA 1470
1021 TATGTGTCTGAACCTGGCTGCTGANTGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1080
1471 TATGTGTCTGAACCTGGCTGCTGANTGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1530
1081 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAATACAG 1140
1531 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAATACAG 1590
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGAGCTGGGCCT 1200
1591 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGAGCTGGGCCT 1650
1201 CCTTCTGATGANTCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG 1260
1651 CCTTCTGATGANTCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG 1710
1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 1301
1711 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 1751

RESULT 3

US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 100.0%; Score 1301; DB 10; Length 3510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCACTTCAGAAGCAGAACCATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
1779 CGACTTTCCAGCACTTCAGAAGCAGAACCATGTACATAGGSCCTTCAAGAGGGAATTGAA 1838

61 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
1839 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 1898
121 GCCTTTGGAAGGACTAGAGAAACTCTTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAG 180
1899 GCCTTTGGAAGGACTAGAGAAACTCTTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAG 1958
181 AGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGGCTGAGGAGGTCAATACTGAGTGGGA 240
1959 AGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGGCTGAGGAGGTCAATACTGAGTGGGA 2018
241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
2019 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2078
301 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
2079 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2138
361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGATCTCTCCAAGATCACCTCGA 420
2139 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGATCTCTCCAAGATCACCTCGA 2198
421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAG 480
2199 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAG 2258
481 TGACCTTGTCTGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCAC 540
2259 TGACCTTGTCTGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCAC 2318
541 TCTGGAAGACCTGAAACACACAGATGGAAGTCTTGAGGTGGCCTCGAGGACCGAGTCAG 600
2319 TCTGGAAGACCTGAAACACACAGATGGAAGTCTTGAGGTGGCCTCGAGGACCGAGTCAG 2378
601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
2379 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2438
661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 720
2439 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 2498
721 CGAGACTCAAACAACCTTGTCTGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
2499 CGAGACTCAAACAACCTTGTCTGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 2558
781 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
2559 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2618
841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCTCAGCTGCATGTGATGCCCTTGACCCAGCA 900
2619 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCTCAGCTGCATGTGATGCCCTTGACCCAGCA 2678
901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATTATTAATTTGTTGACCCAC 960
2679 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATTATTAATTTGTTGACCCAC 2738
961 TATTTATGACCGCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
2739 TATTTATGACCGCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 2798
1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1080
2799 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 2858
1081 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
2859 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 2918
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCCT 1200

Db 2919 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 2978
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTTGGGG 1260
Db 2979 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTTGGGG 3038
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 3039 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 3079

RESULT 4
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-845-416-10
Query Match 100.0%; Score 1301; DB 10; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 1800 CGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 1859
QY 61 AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 1860 AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db 1920 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 1979
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGSCTGAGGAGGTCAATACTAGTGGGA 240
Db 1980 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGSCTGAGGAGGTCAATACTAGTGGGA 2039
QY 241 AAAATTGAACCTGCCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 300
Db 2040 AAAATTGAACCTGCCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 2099
QY 301 CCAGGAACCTTCAAGAGGCCACGGATCAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 2100 CCAGGAACCTTCAAGAGGCCACGGATCAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2159
QY 361 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA 420
Db 2160 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA 2219
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db 2220 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGCCACGTCAA 2279
QY 481 TGACCTTGTCTGCCAGCTTACCACTTTGGGCATTTACCTCTCACCGTATAACCTCAGCAC 540
Db 2280 TGACCTTGTCTGCCAGCTTACCACTTTGGGCATTTACCTCTCACCGTATAACCTCAGCAC 2339

QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2340 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 2399
QY 601 GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 2400 GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2459
QY 661 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 720
Db 2460 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 2519
QY 721 CGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 780
Db 2520 CGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 2579
QY 781 TGACCTGAATAATGTCAATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 2580 TGACCTGAATAATGTCAATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639
QY 841 GAAGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 900
Db 2640 GAAGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 2699
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 960
Db 2700 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 2759
QY 961 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 2760 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 2819
QY 1021 TATGTCTGAACTGGCTGCTGAAATGTTTATGAPACGGGACGAACAGGAGGATCCGTGT 1080
Db 2820 TATGTCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 2879
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 2880 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2939
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 2999
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 3000 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3059
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 3060 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 3100

RESULT 5
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

Query Match 100.0%; Score 1301; DB 10; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
b 2127 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2186

Y 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCA 120
b 2187 AACTAAAGAACTGTAAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCA 2246

Y 121 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGAGAG 180
b 2247 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGAGAG 2306

Y 181 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTAGTGGGA 240
b 2307 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTAGTGGGA 2366

Y 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAACT 300
b 2367 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAACT 2426

Y 301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
b 2427 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2486

Y 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGA 420
b 2487 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGA 2546

Y 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
b 2547 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 2606

Y 481 TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAC 540
b 2607 TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAC 2666

Y 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
b 2667 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 2726

Y 601 GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
b 2727 GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2786

Y 661 TGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
b 2787 TGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2846

Y 721 CGAGACTCAAAACAATTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGC 780
b 2847 CGAGACTCAAAACAATTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGC 2906

Y 781 TGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
b 2907 TGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 2966

Y 841 GAAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 900
b 2967 GAAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 3026

Y 901 CAACCTCAAGCAAAATGACAGAGCCCATGGATATCCTGCAATTAATTAATTTGTTGACCCAC 960
b 3027 CAACCTCAAGCAAAATGACAGAGCCCATGGATATCCTGCAATTAATTAATTTGTTGACCCAC 3086

Y 961 TATTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
b 3087 TATTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 3146

RESULT 6
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match 100.0%; Score 1301; DB 10; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db 3147 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 3206

QY 1081 CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAAGTACAG 1140
Db 3207 CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAAGTACAG 3266

QY 1141 AFACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCCT 1200
Db 3267 AFACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCCT 3326

QY 1201 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTGCATCCTTTGGGG 1260
Db 3327 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTGCATCCTTTGGGG 3386

QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 3387 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 3427

1 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
2268 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2327

61 AACTAAAGAACTGTAAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCA 120
2328 AACTAAAGAACTGTAAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTGACAGAGCA 2387

121 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
2388 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 2447

181 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
2448 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2507

241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAACT 300
2508 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAACT 2567

301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
2568 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2627

361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 420
2628 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 2687

QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAA 480
Db |||||
QY 481 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 541 TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db |||||
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 660
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db |||||
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAAGACTGCA 840
Db |||||
QY 841 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTTCAGCTGCTGATGCTGATGCCCTGGACCAGCA 900
Db |||||
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAAATGTTTGACCCAC 960
Db |||||
QY 961 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db |||||
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db |||||
QY 1081 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAG 1140
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCCT 1200
Db |||||
QY 1201 CCTTCTGCATGATCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db |||||
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db |||||
QY 1301 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 3568
Db |||||

RESULT 7
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 100.0%; Score 1301; DB 10; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 61 AACTAAAGAAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
QY 121 GCCTTTGGAAGGACTAGAGAAACCTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 180
Db |||||
QY 2571 GCCTTTGGAAGGACTAGAGAAACCTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 2630
QY 181 AGCCAGAAATGTCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGCTCAATCTGAGTGGGA 240
Db |||||
QY 241 AAAATTGAAACCTGCACTCCGCTGACTGCGAGAGAAATAATAGATGAGACCCCTTGAAGACT 300
Db |||||
QY 301 CCAGGAACTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTCGCCCAAGCTGAGGTGAT 360
Db |||||
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
Db |||||
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCCTCTGAAAAGAGAACGTGAGCCACGTCAA 480
Db |||||
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 541 TCTGGAAGACCTGAAACACACAGATGGAAGCTTTCGAGGTGGCGTCGAGGACCGAGTCAG 600
Db |||||
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 721 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 780
Db |||||
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
QY 841 GAAGGCCCTTTGCTGGATCTCTTGGCCCTGTTCAGCTGCATGTGATGCCTTGGACCAGCA 900
Db |||||
QY 901 GAAGGCCCTTTGCTGGATCTCTTGGCCCTGTTCAGCTGCATGTGATGCCTTGGACCAGCA 960
Db |||||
QY 961 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db |||||
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db |||||
QY 1081 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAG 1140
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCCT 1200
Db |||||
QY 1201 CCTTCTGCATGATCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db |||||
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db |||||
QY 1301 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 3568
Db |||||

901 CAACCTCAAGCAAAATGACCGCCATGGATATCCCTGCAGATTATTAATTGTTGACCA 960
3351 CAACCTCAAGCAAAATGACCGCCATGGATATCCCTGCAGATTATTAATTGTTGACCA 3410
961 TATTTATGACCGCCTGGAGCAAGACACAACAATTGGTCAACGTCCTCTCTGCGTGA 1020
3411 TATTTATGACCGCCTGGAGCAAGACACAACAATTGGTCAACGTCCTCTCTGCGTGA 3470
1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
3471 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3530
1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 1140
3531 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 3590
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCT 1200
3591 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCT 3650
1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
3651 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3710
1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAAATTG 1301
3711 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAAATTG 3751

RESULT 8
US-09-845-416-31
Sequence 31, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-31

Query Match 100.0%; Score 1301; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
2535 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2594
61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
2595 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2654
121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGAGCTGCCCTCTGAGGAGAG 180
2655 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGAGCTGCCCTCTGAGGAGAG 2714
181 AGCCGAGAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
2715 AGCCGAGAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2774
241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCTTGAAAGACT 300

Db 2775 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 2834
Qy 301 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCRAAGCTGAGGTGAT 360
Db 2835 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCRAAGCTGAGGTGAT 2894
Qy 361 CAAGGGATCCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 2895 CAAGGGATCCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2954
Qy 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAACGTGAGCCACGTCAA 480
Db 2955 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAACGTGAGCCACGTCAA 3014
Qy 481 TGACCTTGTCCCGAGCTTACCATTTTGGGCATTCAGCTCTCAACGTATTAACCTCAGCAC 540
Db 3015 TGACCTTGTCCCGAGCTTACCATTTTGGGCATTCAGCTCTCAACGTATTAACCTCAGCAC 3074
Qy 541 TCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAG 600
Db 3075 TCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAG 3134
Qy 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 3135 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3194
Qy 661 TGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 3195 TGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3254
Qy 721 CGAGACTCAAAACAACCTTGTCTGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 780
Db 3255 CGAGACTCAAAACAACCTTGTCTGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 3314
Qy 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCAATGAAACTCCGAAGACTGCA 840
Db 3315 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3374
Qy 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCCTTGAGCCAGCA 900
Db 3375 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCCTTGAGCCAGCA 3434
Qy 901 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAAGATTATTAATTGTTGACCAC 960
Db 3435 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAAGATTATTAATTGTTGACCAC 3494
Qy 961 TATTTATGACCCCTGGAGCAAGACACACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 3495 TATTTATGACCCCTGGAGCAAGACACACAATTTGGTCAACGTCCTCTCTGCGTGA 3554
Qy 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 3555 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3614
Qy 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db 3615 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 3674
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCT 1200
Db 3675 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCT 3734
Qy 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 3735 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3794
Qy 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
Db 3795 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 3835

Db 3507 GCAGCTGCATGAAGCCACACAGGACCTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 3566
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 3567 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 3626
QY 721 CGAGACTCAACAACTTGTCTGGGACCATCCCAAANTGACAGAGCTCTACCAGTCTTTAGC 780
Db 3627 CGAGACTCAACAACTTGTCTGGGACCATCCCAAANTGACAGAGCTCTACCAGTCTTTAGC 3686
QY 781 TGACCTGAATATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 840
Db 3687 TGACCTGAATATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 3746
QY 841 GAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTGATGCCCTTGGACCAGCA 900
Db 3747 GAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTGATGCCCTTGGACCAGCA 3806
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC 960
Db 3807 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC 3866
QY 961 TATTTATGACCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1020
Db 3867 TATTTATGACCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGGA 3926
QY 1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db 3927 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 3986
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAG 1140
Db 3987 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAG 4046
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 4047 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4106
QY 1201 CCTTCTGCATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 4107 CCTTCTGCATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4166
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 4167 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 4207

RESULT 12
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 100.0%; Score 1301; DB 10; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTTCCAGCAGTTTCCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 60
Db 3025 CGACTTTTCCAGCAGTTTCCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 3084
QY 61 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 120
Db 3085 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 3144
QY 121 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCCAAGAGAGTGCCTCCTGAGGAGAG 180
Db 3145 GCCTTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCCAAGAGAGTGCCTCCTGAGGAGAG 3204
QY 181 AGCCAGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 3205 AGCCAGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3264
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACT 300
Db 3265 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACT 3324
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db 3325 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3384
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 420
Db 3385 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 3444
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACTGAGCCACGTCAC 480
Db 3445 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACTGAGCCACGTCAC 3504
QY 481 TGACCTTGTCTGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 3505 TGACCTTGTCTGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 3564
QY 541 TCTGGAAGACCTGAAACACAGATGGAAGTTTCTGCAGGTGSCCGTTCGAGGACCGAGTCAG 600
Db 3565 TCTGGAAGACCTGAAACACAGATGGAAGTTTCTGCAGGTGSCCGTTCGAGGACCGAGTCAG 3624
QY 601 GCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTC 660
Db 3625 GCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTC 3684
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 3685 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 3744
QY 721 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 3745 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3804
QY 781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3805 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3864
QY 841 GAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCA 900
Db 3865 GAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCA 3924
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGAGATTATTAATTGTTGACCCAC 960
Db 3925 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGAGATTATTAATTGTTGACCCAC 3984
QY 961 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCTCTCTGCGTGGGA 1020
Db 3985 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCTCTCTGCGTGGGA 4044
QY 1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1080
Db 4045 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 4104
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTTGAAGACAAAGTACAG 1140

b 4105 CCTGCTCTTTAAACCTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAGTACAG 4164
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 1200
b 4165 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 4224
1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
b 4225 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4284
1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
b 4285 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 4325

ESULT 13
S-09-845-416-34
Sequence 34, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 34

LENGTH: 4990

TYPE: DNA

ORGANISM: Homo sapiens

S-09-845-416-34

Query Match 100.0%; Score 1301; DB 10; Length 4990;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTCCAGCAGTTTCAAGACGAGAAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAA 60
3049 CGACTTCCAGCAGTTTCAAGACGAGAAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAA 3108
61 AACTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
3109 AACTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 3168
121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 180
3169 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 3228
181 AGCCGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
3229 AGCCGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3288
241 AAAATTGAACCTGCACCTCCGTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
3289 AAAATTGAACCTGCACCTCCGTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3348
301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
3349 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 3408
361 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
3409 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3468
421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
3469 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3528

QY 481 TGACCTTGTCTGCCAGCTTACACATTTGGGCATTTCCAGCTCTCACCGTATATAACCTCAGCAC 540
DB 3529 TGACCTTGTCTGCCAGCTTACACATTTGGGCATTTCCAGCTCTCACCGTATATAACCTCAGCAC 3588
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGGAGTCAAG 600
DB 3589 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGGAGTCAAG 3648
QY 601 GCAGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTC 660
DB 3649 GCAGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTC 3708
QY 661 TGTCCAGGCTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
DB 3709 TGTCCAGGCTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3768
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
DB 3769 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3828
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCCATGAAATCCCGAAGACTGCA 840
DB 3829 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCCATGAAATCCCGAAGACTGCA 3888
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGTGCATGTGATGCCCTTGACCAGCA 900
DB 3889 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGTGCATGTGATGCCCTTGACCAGCA 3948
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
DB 3949 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 4008
QY 961 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1020
DB 4009 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGGA 4068
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
DB 4069 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4128
QY 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTATAAGCACATTTTGGAAAGACAAGTACAG 1140
DB 4129 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTATAAGCACATTTTGGAAAGACAAGTACAG 4188
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCT 1200
DB 4189 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCT 4248
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 4249 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4308
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
DB 4309 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 4349

RESULT 14

US-09-845-416-36

; Sequence 36, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 36

Y 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCTCTCATTGACTCTCTCCAAGATCACCTCGA 420
b |||||
3568 CAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 3627
Y 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480
b |||||
3628 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 3687
Y 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 540
b |||||
3688 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 3747
Y 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCA 600
b |||||
3748 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCA 3807
Y 601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
b |||||
3808 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3867
Y 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720
b |||||
3868 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 3927
Y 721 CCAGACTCAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 780
b |||||
3928 CCAGACTCAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 3987
Y 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCA 840
b |||||
3988 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCA 4047
Y 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCTTGGACCAGCA 900
b |||||
4048 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCTTGGACCAGCA 4107
Y 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 960
b |||||
4108 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 4167
Y 961 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
b |||||
4168 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 4227
Y 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
b |||||
4228 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 4287
Y 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 1140
b |||||
4288 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 4347
Y 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 1200
b |||||
4348 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 4407
Y 1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
b |||||
4408 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4467
Y 1261 CAGTAACATTGAGCCAAAGTGTCGGGAGCTGCTTCCAATTG 1301
b |||||
4468 CAGTAACATTGAGCCAAAGTGTCGGGAGCTGCTTCCAATTG 4508

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 11:55:32 ; Search time 335.709 Seconds
(without alignments)
16463.377 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	1301	100.0	1821	6 AAD37241	Aad37241 Human dys
2	1301	100.0	2169	6 AAD37232	Aad37232 Human dys
3	1301	100.0	3510	6 AAD37240	Aad37240 Human dys
4	1301	100.0	3531	6 AAD37238	Aad37238 Human dys
5	1301	100.0	3858	6 AAD37237	Aad37237 Human dys
6	1301	100.0	3999	6 AAD37234	Aad37234 Human dys
7	1301	100.0	4182	6 AAD37230	Aad37230 Human dys
8	1301	100.0	4476	6 AAD37259	Aad37259 Adeno-ass
9	1301	100.0	4498	6 AAD37258	Aad37258 Adeno-ass
10	1301	100.0	4825	6 AAD37257	Aad37257 Adeno-ass
11	1301	100.0	4848	6 AAD37263	Aad37263 Adeno-ass
12	1301	100.0	4966	6 AAD37256	Aad37256 Adeno-ass
13	1301	100.0	4990	6 AAD37262	Aad37262 Adeno-ass
14	1301	100.0	5060	6 AAD37264	Aad37264 Adeno-ass
15	1301	100.0	5149	6 AAD37255	Aad37255 Adeno-ass
16	1301	100.0	5339	6 ABK81998	Abk81998 DNA encod
17	1301	100.0	5462	6 ABK81999	Abk81999 DNA encod
18	1301	100.0	5952	5 AAD06794	Aad06794 Human dys
19	1301	100.0	8689	6 ABK82000	Abk82000 DNA encod
20	1301	100.0	11058	6 AAD37229	Aad37229 Human dys
21	1301	100.0	11241	6 ABK82005	Abk82005 cDNA enco
22	1301	100.0	11443	6 ABK82002	Abk82002 DNA encod
23	1301	100.0	12923	1 AAN90338	Aan90338 Sequence

24	1301	100.0	13957	6 ABK81959	Abk81959 cDNA enco
25	1301	100.0	13957	6 ABT10904	Abt10904 Human bre
26	1301	100.0	13957	6 ABN95786	Abn95786 Gene #228
27	1301	100.0	13957	6 ABS69900	Abs69900 Human dys
28	1290	99.2	13977	6 ABS70403	Abs70403 Human bon
29	1135.2	87.3	13815	6 ABK81960	Abk81960 cDNA enco
30	1135.2	87.3	13815	6 ABI99799	Abi99799 Mouse isc
31	1135.2	87.3	19307	2 AAT27558	Aat27558 Shuttle v
32	1133.6	87.1	13815	2 AAV18885	Aav18885 Mus muscu
33	1085.2	83.4	4402	3 AAZ48567	Aaz48567 A rod sho
34	1017.6	78.2	3446	6 AAD37242	Aad37242 Human dys
35	1017.6	78.2	4414	6 AAD37260	Aad37260 Adeno-ass
36	1016.6	78.1	5417	6 ABK81997	Abk81997 DNA encod
37	1016	78.1	1434	6 AAD37243	Aad37243 Human dys
38	886	68.1	4402	3 AAZ48568	Aaz48568 A rod sho
39	750.4	57.7	4075	3 AAZ48569	Aaz48569 A rod sho
40	746.2	57.4	3747	3 AAZ48566	Aaz48566 A rod sho
41	692.6	53.2	3163	3 AAZ48571	Aaz48571 A rod sho
42	582.4	44.8	2005	7 ABT33376	Abt33376 NOVX DNA
43	580.8	44.6	2463	7 ABT33375	Abt33375 NOVX DNA
44	570.4	43.8	10705	7 ABT41896	Abt41896 Toxicity
45	567.2	43.6	11096	6 ABK81962	Abk81962 cDNA enco

ALIGNMENTS

RESULT 1					
AAD37241	ID	AAD37241	standard; DNA; 1821 BP.		
XX	AC	AAD37241;			
XX	DT	21-AUG-2002 (first entry)			
XX	DE	Human dystrophin rod, hinge and CR domain regions encoding DNA #2.			
KW	XX	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;			
KW	XX	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;			
KW	XX	Becker muscular dystrophy; ds.			
OS	XX	Homo sapiens.			
XX	PN	WO200183695-A2.			
XX	PD	08-NOV-2001.			
XX	PF	27-APR-2001; 2001WO-US013677.			
XX	PR	28-APR-2000; 2000US-0200777P.			
XX	PA	(XIAO/) XIAO X.			
XX	PI	Xiao X;			
XX	DR	WPI; 2002-049342/06.			
XX	PT	New dystrophin minigene for treating Duchenne or Becker muscular			
XX	PT	dystrophy comprises an N-terminal domain or modified N-terminal domain,			
XX	PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin			
XX	PS	gene.			
XX	PS	Example 1; Page 52-53; 71pp; English.			
XX	CC	The present invention relates to an isolated nucleotide sequence encoding			
XX	CC	a dystrophin minigene. The minigene comprises N-terminal or modified N-			
XX	CC	terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4			
XX	CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The			
XX	CC	invention also relates to a recombinant adeno-associated virus (AAV)			
XX	CC	comprising dystrophin minigene operably linked to an expression control			
XX	CC	element. The dystrophin minigene in operable linkage with an expression			
XX	CC	control element, in a recombinant adeno-associated virus or retrovirus is			
XX	CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular			

CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin DNA fragment encoding rods R23 and R24, hinge H4 and CR domain
CC regions
XX
SQ Sequence 1821 BP; 506 A; 451 C; 447 G; 417 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 1821;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
3b 103 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 162
2Y 61 AACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
3b 163 AACTAAAGAACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCA 222
2Y 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 180
3b 223 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 282
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
3b 283 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 342
QY 241 AAAATTGAACCTGCCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 300
3b 343 AAAATTGAACCTGCCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 402
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
3b 403 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 462
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACTCGA 420
3b 463 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACTCGA 522
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCAGTCAA 480
3b 523 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCAGTCAA 582
QY 481 TGACCTTGCTCGCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 540
3b 583 TGACCTTGCTCGCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 642
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTTGTCAGGTGGCGCTCGAGGACCGAGTCAG 600
3b 643 TCTGGAAGACCTGAACACACAGATGGAAGCTTTGTCAGGTGGCGCTCGAGGACCGAGTCAG 702
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
3b 703 GCAGCTGCATGAAGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTC 762
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
3b 763 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 822
QY 721 CGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
3b 823 CGAGACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 882
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
3b 883 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 942
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGCCCTTGGACAGCA 900
3b 943 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGCCCTTGGACAGCA 1002
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
3b 1003 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 1062

QY 961 TATTTATGACCGCTGGAGCAAGAGCAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
3b 1063 TATTTATGACCGCTGGAGCAAGAGCAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1122
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
3b 1123 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1182
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 1140
3b 1183 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 1242
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCT 1200
3b 1243 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCT 1302
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
3b 1303 CCTTCTGCATGATTTCTATCCAAATTCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1362
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
3b 1363 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1403
RESULT 2
AAD37232
ID AAD37232 standard; DNA; 2169 BP.
XX
AC AAD37232;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin rod, hinge and CR domain regions encoding DNA #1.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 45-46; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human

dystrophin DNA fragment encoding rods R22, R23 and R24, hinge H4 and CR domain regions

Sequence 2169 BP; 623 A; 529 C; 524 G; 493 T; 0 U; 0 Other;

Query Match 100.0%; Score 1301; DB 6; Length 2169;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60

451 CGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 510

61 AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120

511 AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 570

121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGGCCAGAGAGCTGCCCTCTGAGGAGAG 180

571 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGGCCAGAGAGCTGCCCTCTGAGGAGAG 630

181 AGCCCAGAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240

631 AGCCCAGAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 690

241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 300

691 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 750

301 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCRAAGCTGAGGTGAT 360

751 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCRAAGCTGAGGTGAT 810

361 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420

811 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 870

421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTCGAGCCACGTCAA 480

871 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTCGAGCCACGTCAA 930

481 TGACCTTGCTGCCAGCTTACCACTTTTGGGCATTTAGCTCTCACCGTATAACCTCAGCAC 540

931 TGACCTTGCTGCCAGCTTACCACTTTTGGGCATTTAGCTCTCACCGTATAACCTCAGCAC 990

541 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 600

991 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 1050

601 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660

1051 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 1110

661 TGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

1111 TGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1170

721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

1171 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 1230

781 TGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840

1231 TGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1290

841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 900

1291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 1350

901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 960

1351 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 1410

RESULT 3

AAD37240

ID AAD37240 standard; DNA; 3510 BP.

XX

AC AAD37240;

XX

DT 21-AUG-2002 (first entry)

XX

DE Human dystrophin minigene delta3510.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

PT gene.

XX

PS Example 1; Page 51-52; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified N-

CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression

CC control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,

CC	hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC	CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX	
SQ	Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
	Query Match 100.0%; Score 1301; DB 6; Length 3510;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y	1 CGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
Db	1779 CGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 1838
2Y	61 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
Db	1839 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 1898
2Y	121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCCTCCTGAGGAGAG 180
Db	1899 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCCTCCTGAGGAGAG 1958
2Y	181 AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db	1959 AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2018
2Y	241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db	2019 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 2078
QY	301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGTGGCCCAAGCTGAGGTGAT 360
Db	2079 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGTGGCCCAAGCTGAGGTGAT 2138
QY	361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCAATTGACTCTCTCCAAGATCACTCGA 420
Db	2139 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCAATTGACTCTCTCCAAGATCACTCGA 2198
QY	421 GAAAGTCAAGGCATTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db	2199 GAAAGTCAAGGCATTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 2258
QY	481 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db	2259 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 2318
QY	541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 600
Db	2319 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 2378
QY	601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 660
Db	2379 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 2438
QY	661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db	2439 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2498
QY	721 CGAGACTCAAAACAATTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db	2499 CGAGACTCAAAACAATTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 2558
QY	781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db	2559 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2618
QY	841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAGCA 900
Db	2619 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAGCA 2678
QY	901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATAATTGTTTGACCAC 960
Db	2679 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATAATTGTTTGACCAC 2738

QY	961 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db	2739 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 2798
QY	1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db	2799 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 2858
QY	1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db	2859 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2918
QY	1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAGCGCAGGCTGGGCT 1200
Db	2919 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAGCGCAGGCTGGGCT 2978
QY	1201 CCTTCTGCATGATCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db	2979 CCTTCTGCATGATCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3038
QY	1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db	3039 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 3079
RESULT 4	
AAD37238	
ID	AAD37238 standard; DNA; 3531 BP.
XX	
AC	AAD37238;
XX	
DT	21-AUG-2002 (first entry)
XX	
DE	Human dystrophin minigene delta3531.
XX	
KW	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW	Becker muscular dystrophy; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200183695-A2.
XX	
PD	08-NOV-2001.
XX	
PF	27-APR-2001; 2001WO-US013677.
XX	
PR	28-APR-2000; 2000US-0200777P.
XX	
PA	(XIAO/) XIAO X.
XX	
PI	Xiao X;
XX	
DR	WPI; 2002-049342/06.
XX	
PT	New dystrophin minigene for treating Duchenne or Becker muscular
PT	dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT	gene.
XX	
PS	Example 1; Page 50-51; 71pp; English.
XX	
CC	The present invention relates to an isolated nucleotide sequence encoding
CC	a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC	terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC	invention also relates to a recombinant adeno-associated virus (AAV)
CC	comprising dystrophin minigene operably linked to an expression control
CC	element. The dystrophin minigene in operable linkage with an expression
CC	control element, in a recombinant adeno-associated virus or retrovirus is
CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC	dystrophy (BMD) in a mammalian subject. The present sequence is human
CC	dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,

1C hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
2C domain) and 11047-11058 (dystrophin last 3 amino acids)
3X
3Q Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
2b 1800 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 1859
2Y 61 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
2b 1860 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
2Y 121 GCCTTTGGAAGGACTAGAGAACTTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG 180
2b 1920 GCCTTTGGAAGGACTAGAGAACTTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG 1979
2Y 181 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
2b 1980 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2039
2Y 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAAGACT 300
2b 2040 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAAGACT 2099
2Y 301 CCAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 360
2b 2100 CCAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 2159
2Y 361 CAAGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
2b 2160 CAAGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2219
2Y 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
2b 2220 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 2279
2Y 481 TGACCTTGCTCGCCAGCTTTACCACTTTGGGCATTTAGCTCTCACCGTATAACCTCAGCAC 540
2b 2280 TGACCTTGCTCGCCAGCTTTACCACTTTGGGCATTTAGCTCTCACCGTATAACCTCAGCAC 2339
2Y 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGTGGCCGCTCGAGGACCGAGTCAG 600
2b 2340 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGTGGCCGCTCGAGGACCGAGTCAG 2399
2Y 601 GCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
2b 2400 GCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 2459
2Y 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
2b 2460 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2519
2Y 721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCCAGTCTTTAGC 780
2b 2520 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCCAGTCTTTAGC 2579
2Y 781 TGACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCGCATGAACTCCGAAGACTGCA 840
2b 2580 TGACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCGCATGAACTCCGAAGACTGCA 2639
2Y 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTGATGCTTGGACCAGCA 900
2b 2640 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTGATGCTTGGACCAGCA 2699
2Y 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTGACCAC 960
2b 2700 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTGACCAC 2759

QY 961 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db 2760 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGA 2819
QY 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 2820 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 2879
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 2880 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2939
QY 1141 ATACCTTTTCAAGCAAGTGCCAAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGCCT 1200
Db 2940 ATACCTTTTCAAGCAAGTGCCAAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGCCT 2999
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 3000 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3059
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 3060 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 3100
RESULT 5
AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,

CC	hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4	
CC	and CR domain) and 11047-11058 (dystrophin last 3 amino acids)	
XX		
SQ	Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1301; DB 6; Length 3858;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60	
Db		
	2127 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2186	
QY	61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120	
Db		
	2187 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 2246	
QY	121 GCCTTTTGAAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180	
Db		
	2247 GCCTTTTGAAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2306	
QY	181 AGCCCAGAATGTACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240	
Db		
	2307 AGCCCAGAATGTACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2366	
QY	241 AAAATTGAACCTCGCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300	
Db		
	2367 AAAATTGAACCTCGCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2426	
QY	301 CCAGGAACTTCAAGAGSCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360	
Db		
	2427 CCAGGAACTTCAAGAGSCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2486	
QY	361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420	
Db		
	2487 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2546	
QY	421 GAAAGTCAAGGCACITTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480	
Db		
	2547 GAAAGTCAAGGCACITTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 2606	
QY	481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540	
Db		
	2607 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 2666	
QY	541 TCTGGAAGACCTGAACACACCATGGAAGCTTTTGGTCCAGCTCTCGAGGTCGAGGACCGAGTCAG 600	
Db		
	2667 TCTGGAAGACCTGAACACACCATGGAAGCTTTTGGTCCAGCTCTCGAGGTCGAGGACCGAGTCAG 2726	
QY	601 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660	
Db		
	2727 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2786	
QY	661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720	
Db		
	2787 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2846	
QY	721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGC 780	
Db		
	2847 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGC 2906	
QY	781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACTGCA 840	
Db		
	2907 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACTGCA 2966	
QY	841 GAAGGCCCTTTGCTTGGAATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAAGCA 900	
Db		
	2967 GAAGGCCCTTTGCTTGGAATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAAGCA 3026	
QY	901 CAACCTCAAGCAAAATGACAGGCCCATGGATATCTCGAGATTATTAATTGTTTGACCAAC 960	
Db		
	3027 CAACCTCAAGCAAAATGACAGGCCCATGGATATCTCGAGATTATTAATTGTTTGACCAAC 3086	

QY	961 TATTTATGACCGCTGGAGCAAGAGACAAACAAATTTGGTCAACGTCCCTCTCTGCGTGA 1020	
Db		
	3087 TATTTATGACCGCTGGAGCAAGAGACAAACAAATTTGGTCAACGTCCCTCTCTGCGTGA 3146	
QY	1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGACAGGGAGGATCCGTGT 1080	
Db		
	3147 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGACAGGGAGGATCCGTGT 3206	
QY	1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140	
Db		
	3207 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3266	
QY	1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCT 1200	
Db		
	3267 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCT 3326	
QY	1201 CCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTTGGTGAAGTTCATCCTTTGGGG 1260	
Db		
	3327 CCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTTGGTGAAGTTCATCCTTTGGGG 3386	
QY	1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 1301	
Db		
	3387 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 3427	
RESULT 6		
AAD37234		
ID	AAD37234 standard; DNA; 3999 BP.	
XX	AAD37234;	
AC	AAD37234;	
DT	21-AUG-2002 (first entry)	
XX	Human dystrophin minigene delta3990.	
DE	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;	
XX	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;	
KW	Becker muscular dystrophy; ds.	
KW	Homo sapiens.	
XX	WO200183695-A2.	
PN	08-NOV-2001.	
XX	27-APR-2001; 2001WO-US013677.	
PF	28-APR-2000; 2000US-0200777P.	
XX	(XIAO/) XIAO X.	
PA	Xiao X;	
PI	WPI; 2002-049342/06.	
XX	New dystrophin minigene for treating Duchenne or Becker muscular	
PT	dystrophy comprises an N-terminal domain or modified N-terminal domain,	
PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin	
PT	gene.	
XX	Example 1; Page 46-47; 71pp; English.	
PS	The present invention relates to an isolated nucleotide sequence encoding	
XX	a dystrophin minigene. The minigene comprises N-terminal or modified N-	
CC	terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4	
CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The	
CC	invention also relates to a recombinant adeno-associated virus (AAV)	
CC	comprising dystrophin minigene operably linked to an expression control	
CC	element. The dystrophin minigene in operable linkage with an expression	
CC	control element, in a recombinant adeno-associated virus or retrovirus is	
CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular	
CC	dystrophy (BMD) in a mammalian subject. The present sequence is human	
CC	dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,	

XC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
XC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
IX
IQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CGACTTTCCAGCAGTTTCAGAACAGAACGATGTACATAGGGCCCTTCAGAGGGGAATTGAA 60
2268 CGACTTTCCAGCAGTTTCAGAACAGAACGATGTACATAGGGCCCTTCAGAGGGGAATTGAA 2327
61 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
2328 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 2387
121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCCTCCTGAGGAGAG 180
2388 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCCTCCTGAGGAGAG 2447
181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
2448 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2507
241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTTGAAAGACT 300
2508 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTTGAAAGACT 2567
301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
2568 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2627
361 CAAGGATCTTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGA 420
2628 CAAGGATCTTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGA 2687
421 GAAAGTCAAGGCACCTTCGAGAGAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA 480
2688 GAAAGTCAAGGCACCTTCGAGAGAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA 2747
481 TGACCTTGCTGCCAGCTTACACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 540
2748 TGACCTTGCTGCCAGCTTACACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 2807
541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCGTCGAGACCCGAGTCAG 600
2808 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCGTCGAGACCCGAGTCAG 2867
601 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
2868 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 2927
661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720
2928 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 2987
721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 780
2988 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 3047
781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
3048 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3107
841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCA 900
3108 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCA 3167
901 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTCGAGATTATTAAATTGTTGACCAC 960
3168 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTCGAGATTATTAAATTGTTGACCAC 3227

QY 961 TATTATGACCGCCTGGAGCAAGACAGACACAATTTGGTCAACGTCCCTCTCTCGGTGGA 1020
Db 3228 TATTATGACCGCCTGGAGCAAGACAGACACAATTTGGTCAACGTCCCTCTCTCGGTGGA 3287
QY 1021 TATGTCTCTGAACCTGGCTGCTGATGTTTATGATACGGACGAACAGGGAGGATCCGTGT 1080
Db 3288 TATGTCTCTGAACCTGGCTGCTGATGTTTATGATACGGACGAACAGGGAGGATCCGTGT 3347
QY 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAACACATTTTGGAAAGACAAGTACAG 1140
Db 3348 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAACACATTTTGGAAAGACAAGTACAG 3407
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAGCGCAGGCTGGGCCT 1200
Db 3408 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAGCGCAGGCTGGGCCT 3467
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 3468 CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3527
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 3528 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 3568
RESULT 7
AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 43-44; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,

CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCGAGTTCAGAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAATTGAA 60
Db 2451 CGACTTTCCAGCGAGTTCAGAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAATTGAA 2510
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2511 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2570
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 2571 GCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 2630
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 240
Db 2631 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 2690
QY 241 AAAATTGAACCTGCCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 300
Db 2691 AAAATTGAACCTGCCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 2750
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db 2751 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2810
QY 361 CAAGGGATCCTGGCAGCCGCTGGCGATCTCTCATTTGACTCTCTCCAAGATCACTCGA 420
Db 2811 CAAGGGATCCTGGCAGCCGCTGGCGATCTCTCATTTGACTCTCTCCAAGATCACTCGA 2870
QY 421 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db 2871 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 2930
QY 481 TGACCTTGCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 2931 TGACCTTGCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 2990
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2991 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3050
QY 601 GCAGCTGCATGAAGCCACAGGGAACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
Db 3051 GCAGCTGCATGAAGCCACAGGGAACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 3110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 3111 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3170
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 3171 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3230
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3231 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCA 3290
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGCA 900
Db 3291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGCA 3350
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCAC 960
Db 3351 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCAC 3410

QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db 3411 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 3470
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCGTGT 1080
Db 3471 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCGTGT 3530
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAAGACAAGTACAG 1140
Db 3531 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAAGACAAGTACAG 3590
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGACACGCGCAGGCTGGGCT 1200
Db 3591 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGACACGCGCAGGCTGGGCT 3650
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 3651 CCTTCTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3710
QY 1261 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db 3711 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 3751
RESULT 8
AAD37259
ID AAD37259 standard; DNA; 4476 BP.
XX
AC AAD37259;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN W0200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 63-65; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

Query Match 100.0%; Score 1301; DB 6; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
|||||
2535 CGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2594
|||||

61 AACTAAGAACTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
|||||
2595 AACTAAGAACTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 2654
|||||

121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
|||||
2655 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 2714
|||||

181 AGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
|||||
2715 AGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2774
|||||

241 AAAATTGAACCTGCACCTCCGCTGACTGCGGAGAGAGAAATATAGATGAGACCCCTTGAAAGACT 300
|||||
2775 AAAATTGAACCTGCACCTCCGCTGACTGCGGAGAGAGAAATATAGATGAGACCCCTTGAAAGACT 2834
|||||

301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
|||||
2835 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2894
|||||

361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
|||||
2895 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2954
|||||

421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCAGCTCAA 480
|||||
2955 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCAGCTCAA 3014
|||||

481 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATACCTCAGCAC 540
|||||
3015 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATACCTCAGCAC 3074
|||||

541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGCTGCGCGTGGAGCCGAGTCA 600
|||||
3075 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGCTGCGCGTGGAGCCGAGTCA 3134
|||||

601 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTTCCAGTC 660
|||||
3135 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTTCCAGTC 3194
|||||

661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
|||||
3195 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3254
|||||

721 CGAGACTCAAACTGCTGGGAGAGCCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 780
|||||
3255 CGAGACTCAAACTGCTGGGAGAGCCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 3314
|||||

781 TGACCTGAATAATGTGATGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
|||||
3315 TGACCTGAATAATGTGATGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3374
|||||

841 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTTGGACAGCA 900
|||||
3375 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTTGGACAGCA 3434
|||||

901 CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTAATTAATTTGTTGACCAC 960
|||||
3435 CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTAATTAATTTGTTGACCAC 3494
|||||

961 TATTTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
|||||
3495 TATTTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA 3554
|||||

1021 TATGTGTCTGAACCTGGCTGCTGATGTTTATGATACCGGACGAACAGGAGGATCCGTGT 1080
|||||
3555 TATGTGTCTGAACCTGGCTGCTGATGTTTATGATACCGGACGAACAGGAGGATCCGTGT 3614
|||||

1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 1140
|||||
3615 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 3674
|||||

1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGCCT 1200
|||||
3675 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGCCT 3734
|||||

1201 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
|||||
3735 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3794
|||||

1261 CAGTAACATTGAGCAAGTGTCCGGAGCTCTTCCAATTTG 1301
|||||
3795 CAGTAACATTGAGCAAGTGTCCGGAGCTCTTCCAATTTG 3835
|||||

RESULT 9
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX
AC AAD37258;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 62-63; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2557 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2616
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2617 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2677 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 2736
QY 181 AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2737 AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2796
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2856
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGAAT 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2857 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGAAT 2916
QY 361 CAAGGATCCTGGCAGCCGCTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2917 CAAGGATCCTGGCAGCCGCTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 2976
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2977 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 3036
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3037 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3096
QY 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAG 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3097 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAG 3156
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTTGTTCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3157 GCAGCTGCATGAAGCCACAGGGACTTTTGTTCAGCATCTCAGCACTTTCTTTCCACGTC 3216
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3217 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3276
QY 721 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3277 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3336
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3337 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 3396
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA 900
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3397 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA 3456
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 3457 CAACCTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 3516
QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGA 1020
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3517 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGA 3576
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 1080
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3577 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 3636
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3637 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3696
QY 1141 ATACTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 1200
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3697 ATACTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 3756
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3757 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3816
QY 1261 CAGTAACATTTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3817 CAGTAACATTTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 3857
RESULT 10
AAD37257
ID AAD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
FI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular.
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 61-62; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression

CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 1 CGACTTTCAGAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
DB 2884 CGACTTTCAGAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2943
2Y 61 AACTAAAGAACTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 2944 AACTAAAGAACTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3003
2Y 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 3004 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 3063
2Y 181 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGA 240
DB 3064 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGA 3123
2Y 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
DB 3124 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 3183
2Y 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 360
DB 3184 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 3243
2Y 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 420
DB 3244 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 3303
2Y 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCCAGCTCAA 480
DB 3304 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCCAGCTCAA 3363
2Y 481 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 3364 TGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3423
2Y 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 600
DB 3424 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 3483
2Y 601 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 660
DB 3484 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 3543
2Y 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
DB 3544 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3603
2Y 721 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
DB 3604 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3663
2Y 781 TGACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB 3664 TGACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3723
2Y 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGATGCTGATGCTTGGACGCA 900
DB 3724 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGATGCTGATGCTTGGACGCA 3783
2Y 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATATTATTGTTTGACCCAC 960

DB 3784 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTATTGTTGACCCAC 3843
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
DB 3844 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGA 3903
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 1080
DB 3904 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 3963
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTCTGTAAAGCACATTTGGAAGACAAGTACAG 1140
DB 3964 CCTGTCTTTTAAACTGGCATCATTTCCCTCTGTAAAGCACATTTGGAAGACAAGTACAG 4023
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 1200
DB 4024 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 4083
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTCATCCTTTGGGGG 1260
DB 4084 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTCATCCTTTGGGGG 4143
QY 1261 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 1301
DB 4144 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG 4184
RESULT 11
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
QY 2967 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3026
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 180
Db |||||
QY 3027 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 3086
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATACTAGTGGGA 240
Db |||||
QY 3087 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATACTAGTGGGA 3146
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 300
Db |||||
QY 3147 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 3206
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db |||||
QY 3207 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACTCAAGCTGCGCCCAAGCTGAGGTGAT 3266
QY 361 CAAGGATCCTGGCAGCCCCGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACTCGA 420
Db |||||
QY 3267 CAAGGATCCTGGCAGCCCCGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACTCGA 3326
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db |||||
QY 3327 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 3386
QY 481 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 3387 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3446
QY 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGACCGAGTCAG 600
Db |||||
QY 3447 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGACCGAGTCAG 3506
QY 601 GCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
Db |||||
QY 3507 GCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 3566
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 3567 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3626
QY 721 CGAGACTCAAAACAACTTGTGCGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGC 780
Db |||||
QY 3627 CGAGACTCAAAACAACTTGTGCGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGC 3686
QY 781 TGACCTGAATAATGTCAAGTCTCAGCTTATAGGAGTCCCATGAAATCCCGAAGACTGCA 840
Db |||||
QY 3687 TGACCTGAATAATGTCAAGTCTCAGCTTATAGGAGTCCCATGAAATCCCGAAGACTGCA 3746
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 900
Db |||||
QY 3747 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCAGCA 3806

QY 901 CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCAC 960
Db |||||
QY 3807 CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAAATTGTTGACCAC 3866
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTCGTGGGA 1020
Db |||||
QY 3867 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTCGTGGGA 3926
QY 1021 TATGTGCTGAAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db |||||
QY 3927 TATGTGCTGAAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3986
QY 1081 CCTGTCTTTTAAAACTGGCATCATCTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db |||||
QY 3987 CCTGTCTTTTAAAACTGGCATCATCTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4046
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCT 1200
Db |||||
QY 4047 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCT 4106
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 1260
Db |||||
QY 4107 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 4166
QY 1261 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
Db |||||
QY 4167 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 4207

RESULT 12

AAD37256

ID AAD37256 standard; DNA; 4966 BP.

XX

AC AAD37256;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.

OS Unidentified.

OS Chimeric.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

PT gene.

XX

PS Example 1; Page 59-60; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified N-

CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX

SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4966;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTGAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
DB 3025 CGACTTTCCAGCAGTTGAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3084
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 3085 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3144
QY 121 GCCTTTGGAAGGACTAGAGAAAACCTCTACCAAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
DB 3145 GCCTTTGGAAGGACTAGAGAAAACCTCTACCAAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 3204
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGAGGGA 240
DB 3205 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGAGGGA 3264
QY 241 AAAATTGAACCTGCACCTCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTTGAAAGACT 300
DB 3265 AAAATTGAACCTGCACCTCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTTGAAAGACT 3324
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
DB 3325 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 3384
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
DB 3385 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3444
QY 421 GAAAGTCAAGGCACCTCGAGGAGAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
DB 3445 GAAAGTCAAGGCACCTCGAGGAGAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 3504
QY 481 TGACCTTGCTGGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTATAACCTCAGCAC 540
DB 3505 TGACCTTGCTGGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTATAACCTCAGCAC 3564
QY 541 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
DB 3565 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3624
QY 601 GCAGCTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
DB 3625 GCAGCTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3684
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720
DB 3685 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 3744
QY 721 CGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
DB 3745 CGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 3804
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB 3805 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3864
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCTTGGACCAAGCA 900

DB 3865 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGACCAGCA 3924
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATFCTCTGAGATTATTAAATGTTTACCAC 960
DB 3925 CAACCTCAAGCAAAATGACAGCCCATGGATFCTCTGAGATTATTAAATGTTTACCAC 3984
QY 961 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
DB 3985 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 4044
QY 1021 TATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
DB 4045 TATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 4104
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 1140
DB 4105 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 4164
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 1200
DB 4165 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 4224
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 4225 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4284
QY 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
DB 4285 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 4325

RESULT 13

AAD37262

ID AAD37262 standard; DNA; 4990 BP.

XX AAD37262;

AC AAD37262;

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.

OS Cytomegalovirus.

OS Unidentified.

OS Chimeric.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

PT gene.

XX

PS Example 1; Page 67-68; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified N-

CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;
Query Match 100.0%; Score 1301; DB 6; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTTCAGAAAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 61 AACTAAAGAACTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
QY 3109 AACTAAAGAACTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 3168
QY 121 GCCTTTGGAAGACTAGAGAACTCTTACAGGAGCCCAAGAGAGCTGCTCTCTGAGGAGAG 180
Db |||||
QY 3169 GCCTTTGGAAGACTAGAGAACTCTTACAGGAGCCCAAGAGAGCTGCTCTCTGAGGAGAG 3228
QY 181 AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db |||||
QY 3229 AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3288
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGGCAGAGAAAATAGATGAGACCCCTTCAAAGACT 300
Db |||||
QY 3289 AAAATTGAACCTGCACCTCCGCTGACTGGGCAGAGAAAATAGATGAGACCCCTTCAAAGACT 3348
QY 301 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTCAGGTGAT 360
Db |||||
QY 3349 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTCAGGTGAT 3408
QY 361 CAAGGGATCTCGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db |||||
QY 3409 CAAGGGATCTCGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3468
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db |||||
QY 3469 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCCTCTGAAAGAGAACTGAGCCACGTCAA 3528
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 3529 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3588
QY 541 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGCGCGTCGAGGACCGAGTCAG 600
Db |||||
QY 3589 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGCGCGTCGAGGACCGAGTCAG 3648
QY 601 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db |||||
QY 3649 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3708
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTGCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db |||||
QY 3709 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTGCGCCAAACAAAGTGCCTACTATATCAACCA 3768
QY 721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db |||||
QY 3769 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3828
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGGAAGACTGCA 840
Db |||||
QY 3829 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGGAAGACTGCA 3888

QY 841 GAAGGCCCTTGTGGATCTCTTGAGCCCTGTGACCTGTCATGTGATGCTTGGACCAGCA 900
Db |||||
QY 3889 GAAGGCCCTTGTGGATCTCTTGAGCCCTGTGACCTGTCATGTGATGCTTGGACCAGCA 3948
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCCTCAGATTATTAATTGTTGACCAC 960
Db |||||
QY 3949 CAACCTCAAGCAAAATGACAGCCCATGGATATCCCTCAGATTATTAATTGTTGACCAC 4008
QY 961 TATTTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 1020
Db |||||
QY 4009 TATTTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 4068
QY 1021 TATGTGTCTGAACCTGGCTGCTGATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db |||||
QY 4069 TATGTGTCTGAACCTGGCTGCTGATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 4128
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 1140
Db |||||
QY 4129 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 4188
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGAGGTGGGCCT 1200
Db |||||
QY 4189 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGAGGTGGGCCT 4248
QY 1201 CCTTCTGCATGATCTATCCAAATTCAGAGACAGTTGGTGAAGTTGCATCCCTTGGGGG 1260
Db |||||
QY 4249 CCTTCTGCATGATCTATCCAAATTCAGAGACAGTTGGTGAAGTTGCATCCCTTGGGGG 4308
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTG 1301
Db |||||
QY 4309 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTG 4349
RESULT 14
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX
AC AAD37264;
XX DT 21-AUG-2002 (first entry)
XX DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX PN WO200183695-A2.
XX PN 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US013677.
XX PR 28-APR-2000; 2000US-0200777P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 70-71; 71pp; English.
XX

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene operably linked with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a small polyA signal sequence

Sequence 5060 BP: 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Query Match 100.0%; Score 1301; DB 6; Length 5060;

Query Match
Best Local Similarity
Best Local Similarity
100.0%;
100.0%;
Pred. No. 0;
Pred. No. 0;

BEST LOCAL SIMILARITY 100.0%, req. NO. 0;
Matches 1301: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

2Y	1	CGACTTTCCAGCAGTTT	CAGAAGCAGAA	CGATGTACATAGGGCCTT	CAAGAGGGAATTGAA	60
3b	3119	CGACTTTCCAGCAGTTT	CAGAAGCAGAA	CGATGTACATAGGGCCTT	CAAGAGGGAATTGAA	3178
2Y	61	AACATAAGAACCTGTAAT	CATGAGTACTCTTTGAGACTGT	TACGAATATTTCTGCACAGACA	120	
3b	3179	AACATAAGAACCTGTAAT	CATGAGTACTCTTTGAGACTGT	TACGAATATTTCTGCACAGACA	3238	
2Y	121	GCCTTTTGAAGGACTAGAGAA	ACTCTACCCAGGAGCCCAGAGAGCTG	CCCTCCTGAGGAGAG	180	
3b	3239	GCCTTTTGAAGGACTAGAGAA	ACTCTACCCAGGAGCCCAGAGAGCTG	CCCTCCTGAGGAGAG	3298	
2Y	181	AGCCAGAAATGTCACTCGGCTT	CTACGAAAGCAGGCTGAGGAGGTCA	ATACTAGTGGGA	240	
3b	3299	AGCCAGAAATGTCACTCGGCTT	CTACGAAAGCAGGCTGAGGAGGTCA	ATACTAGTGGGA	3358	
2Y	241	AAAATTGAACCTGCACTCCGCTG	ACTGGCAGAGAAAATAGATGAGACCC	TTTGAAGACT	300	
3b	3359	AAAATTGAACCTGCACTCCGCTG	ACTGGCAGAGAAAATAGATGAGACCC	TTTGAAGACT	3418	
2Y	301	CCAGGAACTTCAAGAGGCCACGGAT	GAGCTGGACCTCAAGCTGCGCCAA	GGCTGAGGTGAT	360	
3b	3419	CCAGGAACTTCAAGAGGCCACGGAT	GAGCTGGACCTCAAGCTGCGCCAA	GGCTGAGGTGAT	3478	
2Y	361	CAAGGGATCTTGGCAGCCCGTGGCG	ATCTCTCATTTGACTCTCTCCAAGAT	CACCTCGA	420	
3b	3479	CAAGGGATCTTGGCAGCCCGTGGCG	ATCTCTCATTTGACTCTCTCCAAGAT	CACCTCGA	3538	
2Y	421	GAAAGTCAAGGCACTTCGAGGAGAAA	ATTGCGCCTCTGAAAGAGAA	CGTGAGCCACGTCAA	480	
3b	3539	GAAAGTCAAGGCACTTCGAGGAGAAA	ATTGCGCCTCTGAAAGAGAA	CGTGAGCCACGTCAA	3598	
2Y	481	TGACCTTGCTCGCCAGCTTACCAC	TTTGGGCACTTCAGCTCTCACCGTATA	ACCTCAGCAC	540	
3b	3599	TGACCTTGCTCGCCAGCTTACCAC	TTTGGGCACTTCAGCTCTCACCGTATA	ACCTCAGCAC	3658	
2Y	541	TCTGGAAGACCTGAACAC	CAGATGGAAGCTTCTGCAAGTGGCCGT	CGAGGACCGAGTCAG	600	
3b	3659	TCTGGAAGACCTGAACAC	CAGATGGAAGCTTCTGCAAGTGGCCGT	CGAGGACCGAGTCAG	3718	
2Y	601	GCAGCTGCATGAAGCCCA	CAGGGACTTTTGGTCCAGCATCTCAGCACT	TTTCTTTCCACGTC	660	
3b	3719	GCAGCTGCATGAAGCCCA	CAGGGACTTTTGGTCCAGCATCTCAGCACT	TTTCTTTCCACGTC	3778	
2Y	661	TGTCCAGGGTCCCTGGGAGAGAGCC	ATCTCGCCAAACAAAGTGCCTTACTATAT	CAACCA	720	
3b	3779	TGTCCAGGGTCCCTGGGAGAGAGCC	ATCTCGCCAAACAAAGTGCCTTACTATAT	CAACCA	3838	
2Y	721	CGAGACTCAAAACA	ACTTGCTGGGACCATCCCAAAATGACAGAGCT	CTACCAGTCTTTAGC	780	
3b	3839	CGAGACTCAAAACA	ACTTGCTGGGACCATCCCAAAATGACAGAGCT	CTACCAGTCTTTAGC	3898	
2Y	781	TGACCTGAATAAT	TGTCAGATTCTCAGCTTATAGGACTG	CGCATGAAACTCCGAGAGCTGCA	840	

Example 1; Page 57-59; 71pp; English.

PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (CK) promoter and a small polyA signal sequence
XX

SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match 100.0%; Score 1301; DB 6; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA	60
DB	3208	CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA	3267
QY	61	AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	120
DB	3268	AACTAAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	3327
QY	121	GCCTTTGAAGGACTAGAGAAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAG	180
DB	3328	GCCTTTGAAGGACTAGAGAAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAG	3387
QY	181	AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGA	240
DB	3388	AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGA	3447
QY	241	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAAGACT	300
DB	3448	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAAGACT	3507
QY	301	CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCACAGCTGAGGTGAT	360
DB	3508	CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCACAGCTGAGGTGAT	3567
QY	361	CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA	420
DB	3568	CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA	3627
QY	421	GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA	480
DB	3628	GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA	3687
QY	481	TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTATTAACCTCAGCAC	540
DB	3688	TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTATTAACCTCAGCAC	3747
QY	541	TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAG	600
DB	3748	TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAG	3807
QY	601	GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC	660
DB	3808	GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC	3867
QY	661	TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA	720
DB	3868	TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA	3927
QY	721	CGAGACTCAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGC	780
DB	3928	CGAGACTCAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGC	3987

QY	781	TGACCTGAATAATGTGATGCTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA	840
DB	3988	TGACCTGAATAATGTGATGCTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA	4047
QY	841	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTGATGCTTGGACCAAGCA	900
DB	4048	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTGATGCTTGGACCAAGCA	4107
QY	901	CAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC	960
DB	4108	CAACTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC	4167
QY	961	TATTTATGACCGCTTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGGA	1020
DB	4168	TATTTATGACCGCTTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGGA	4227
QY	1021	TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGCTGT	1080
DB	4228	TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGCTGT	4287
QY	1081	CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG	1140
DB	4288	CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG	4347
QY	1141	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCCAGCGGAGGCTGGGCCT	1200
DB	4348	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCCAGCGGAGGCTGGGCCT	4407
QY	1201	CCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCAATCCTTTGGGGG	1260
DB	4408	CCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCAATCCTTTGGGGG	4467
QY	1261	CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG	1301
DB	4468	CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG	4508

Search completed: April 4, 2004, 18:20:19
Job time : 338.709 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

MM nucleic - nucleic search, using sw model
Run on: April 4, 2004, 12:04:07 ; Search time 3391.8 Seconds
(without alignments)
16625.161 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1301	100.0	5339	6	AX538620	AX538620 Sequence
2	1301	100.0	5462	6	AX538621	AX538621 Sequence
3	1301	100.0	5952	6	AR304538	AR304538 Sequence
4	1301	100.0	5952	6	AX114289	AX114289 Sequence
5	1301	100.0	8689	6	AX538622	AX538622 Sequence
6	1301	100.0	11443	6	AX538624	AX538624 Sequence
7	1301	100.0	12057	6	AX538627	AX538627 Sequence
8	1301	100.0	12446	9	HSDMDR	X14298 Human mRNA
9	1301	100.0	13957	6	AX409637	AX409637 Sequence
10	1301	100.0	13957	6	AX538581	AX538581 Sequence
11	1301	100.0	13957	9	HUMDYS	M18533 Homo sapien
12	1290	99.2	13977	6	AR220819	AR220819 Sequence
13	1154.4	88.7	13887	4	AF070485	AF070485 Canis fam
14	1135.2	87.3	13815	6	AX306153	AX306153 Sequence
15	1135.2	87.3	13815	6	AX538582	AX538582 Sequence
16	1135.2	87.3	13815	10	MUSDYSA	M68859 Mouse dyster
17	1135.2	87.3	19307	6	AR093392	AR093392 Sequence
18	1135.2	87.3	19307	6	AR142592	AR142592 Sequence
19	1085.2	83.4	4402	6	E30219	E30219 Shortened d
20	1016.6	78.1	5417	6	AX538619	AX538619 Sequence
21	904.8	69.5	13575	5	GGDYS	X13369 Chicken mRN
22	886	68.1	4402	6	E30220	E30220 Shortened d
23	750.4	57.7	4075	6	E30221	E30221 Shortened d
24	746.2	57.4	3747	6	E30218	E30218 Shortened d
25	722.4	55.5	2654	5	FSCDYSTRO	M37645 Torpedo cal
26	692.6	53.2	3163	6	E30223	E30223 Shortened d
27	608.2	46.7	3521	5	AF339031	AF339031 Danio rer
28	584	44.9	2110	9	HUMDMDX	M92650 Human Duche
29	584	44.9	4658	9	BC028720	BC028720 Homo sapi
30	582.4	44.8	2005	6	AX817318	AX817318 Sequence
31	580.8	44.6	2463	6	AX817316	AX817316 Sequence
32	570.4	43.8	10705	10	RNAJ2967	AJ002967 Rattus no
33	567.2	43.6	3161	10	NMGUTRPH	X83506 M.musculus
34	567.2	43.6	11096	6	AX538584	AX538584 Sequence
35	567.2	43.6	11096	10	MMY12229	Y12229 M.musculus
36	553.8	42.6	6045	6	A63605	A63605 Sequence 7
37	553.8	42.6	6045	6	AR281528	AR281528 Sequence
38	553.8	42.6	6059	6	AX107972	AX107972 Sequence
39	553.8	42.6	10302	6	AX538583	AX538583 Sequence
40	553.8	42.6	10302	9	HSMUPS	X69086 H.sapiens m
41	553.8	42.6	10320	6	A63607	A63607 Sequence 9
42	553.8	42.6	10320	6	AR281529	AR281529 Sequence
43	551	42.4	3499	9	HSU43519	U43519 Human dyster
44	551	42.4	5106	6	AX552248	AX552248 Sequence
45	546.8	42.0	3073	10	AF195788	AF195788 Rattus no

ALIGNMENTS

RESULT 1	AX538620	AX538620	5339 bp	DNA	linear	PAT 23-NOV-2002
LOCUS	AX538620	Sequence 40 from Patent WO0229056.				
DEFINITION	AX538620					
ACCESSION	AX538620					
VERSION	AX538620.1	GI:25271166				
KEYWORDS						
SOURCE		synthetic construct				
ORGANISM		synthetic construct				
		artificial sequences.				
REFERENCE	1					
AUTHORS		Chamberlain, J.S. and Harper, S.Q.				
TITLE		Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL		Patent: WO 0229056-A 40 11-APR-2002;				
		THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
----------	--	---------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

db 2356 AAAATTGAACCTGCACCTCCGCTGACTCGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2415
2y 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
db 2416 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 2475
2y 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACTCGA 420
db 2476 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCTCAAGATCACTCGA 2535
2y 421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
db 2536 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 2595
2y 481 TGACCTTGCTCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
db 2596 TGACCTTGCTCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 2655
2y 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 600
db 2656 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAG 2715
2y 601 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
db 2716 GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 2775
2y 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
db 2776 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2835
2y 721 CGAGACTCAAAACAACTTGTGGACCAATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 780
db 2836 CGAGACTCAAAACAACTTGTGGACCAATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 2895
2y 781 TGACCTGAATATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
db 2896 TGACCTGAATATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2955
2y 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGACAGCA 900
db 2956 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGACAGCA 3015
2y 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTAATTAATTTTGACCAAC 960
db 3016 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTAATTAATTTTGACCAAC 3075
2y 961 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAAGTCCCTCTCTGCGTGA 1020
db 3076 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAAGTCCCTCTCTGCGTGA 3135
2y 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGTCCGTGT 1080
db 3136 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGTCCGTGT 3195
2y 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
db 3196 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 3255
2y 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
db 3256 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 3315
2y 1201 CCTTCTGCATGATTTATCCAAATTTCCAAAGACAGTTGGGTGAAGTGCATCCTTTGGGG 1260
db 3316 CCTTCTGCATGATTTATCCAAATTTCCAAAGACAGTTGGGTGAAGTGCATCCTTTGGGG 3375
2y 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
db 3376 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 3416

LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. .5952
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 5952;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTCCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAA 60
DB 3402 CGACTTTCAGCAGTTCCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAA 3461
QY 61 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATAATTTCTGACAGAGCA 120
DB 3462 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATAATTTCTGACAGAGCA 3521
QY 121 GCCTTTTGGAAAGGACTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 3522 GCCTTTTGGAAAGGACTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 3581
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGAGCTGAGGAGTCAATACTGAGTGGGA 240
DB 3582 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGAGCTGAGGAGTCAATACTGAGTGGGA 3641
QY 241 AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 300
DB 3642 AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 3701
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT 360
DB 3702 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT 3761
QY 361 CAAGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
DB 3762 CAAGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3821
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
DB 3822 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3881
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 540
DB 3882 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 3941
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
DB 3942 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 4001
QY 601 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACCTTTTCCACGTC 660
DB 4002 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACCTTTTCCACGTC 4061
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
DB 4062 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 4121
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 4122 CGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 4181

QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840

Db 4182 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 4241

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCAGCA 900

Db 4242 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCAGCA 4301

QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGAGATTATTAATTTGTGACCAC 960

Db 4302 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGAGATTATTAATTTGTGACCAC 4361

QY 961 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020

Db 4362 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 4421

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 1080

Db 4422 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 4481

QY 1081 CCTGTCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140

Db 4482 CCTGTCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 4541

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGCGCAGGCTGGGCT 1200

Db 4542 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGCGCAGGCTGGGCT 4601

QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAGACAGTTGGGTGAAGTGCATCCTTTGGGG 1260

Db 4602 CCTTCTGCATGATTCTATCCAAATTTCCAGACAGTTGGGTGAAGTGCATCCTTTGGGG 4661

QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301

Db 4662 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 4702

RESULT 4

AX114289

LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001

DEFINITION Sequence 1 from Patent WO0129243.

ACCESSION AX114289

VERSION AX114289.1 GI:14031259

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Paul,X.L. and Xiao,X.

TITLE Method and vector for producing and transferring trans -spliced peptides

JOURNAL Patent: WO 0129243-A 1 26-APR-2001; DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)

FEATURES

source

1. .5952

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

misc_feature 2897..2898

/note="S4 junction site"

misc_feature 3198..3199

/note="S2 junction site"

ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 5952;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60

Db 3402 CGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3461

QY 61 AACTAAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120

Db 3462 AACTAAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3521

QY 121 GCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGAGGCCAGAGAGCTGCTCCTGAGGAGAG 180

Db 3522 GCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGAGGCCAGAGAGCTGCTCCTGAGGAGAG 3581

QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240

Db 3582 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 3641

QY 241 AAAATTGAACCTGCACTCGGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGACT 300

Db 3642 AAAATTGAACCTGCACTCGGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGACT 3701

QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360

Db 3702 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 3761

QY 361 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420

Db 3762 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3821

QY 421 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGGCTCTGAAAGAGAACGTGAGCCACGTCAA 480

Db 3822 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGGCTCTGAAAGAGAACGTGAGCCACGTCAA 3881

QY 481 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCCGTATAACCTCAGCAC 540

Db 3882 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCCGTATAACCTCAGCAC 3941

QY 541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600

Db 3942 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 4001

QY 601 GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC 660

Db 4002 GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC 4061

QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720

Db 4062 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 4121

QY 721 CGAGACTCAAAACAACTTGTGTTGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780

Db 4122 CGAGACTCAAAACAACTTGTGTTGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 4181

QY 781 TGACCTGAATAATGTGATTTCTCAGATTCTTAGGACTGCCATGAACTCCGAAGACTGCA 840

Db 4182 TGACCTGAATAATGTGATTTCTCAGATTCTTAGGACTGCCATGAACTCCGAAGACTGCA 4241

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACCAGCA 900

Db 4242 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCCTGTGATGCTTGGACCAGCA 4301

QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGAGATTATTAATTTGTGACCAC 960

Db 4302 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGAGATTATTAATTTGTGACCAC 4361

QY 961 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020

Db 4362 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 4421

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080

Db 4422 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 4481

QY 1081 CCTGTCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140

4482 CCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4541
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGCCT 1200
4542 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGCCT 4601
1201 CCTTCTGATGATTTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGTCATCCTTTGGGG 1260
4602 CCTTCTGATGATTTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGTCATCCTTTGGGG 4661
1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 1301
4662 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 4702

RESULT 5
AX538622 8689 bp DNA linear PAT 23-NOV-2002
LOCUS Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .8689
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 8689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CGACTTTCCAGAGTTTCAAGACGACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
3448 CGACTTTCCAGAGTTTCAAGACGACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 3507
61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
3508 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3567
121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 180
3568 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCCTGAGGAGAG 3627
181 AGCCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
3628 AGCCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3687
241 AAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAATAGATGAGACCCCTTGAAAGACT 300
3688 AAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAATAGATGAGACCCCTTGAAAGACT 3747
301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
3748 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 3807
361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCTCCAAGATCACCTCGA 420
3808 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCTCCAAGATCACCTCGA 3867
421 GAAAGTCAAGGCACCTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
3868 GAAAGTCAAGGCACCTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3927

QY 481 TGACCTTGTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTACCGTATAAACCTCAGCAC 540
Db 3928 TGACCTTGTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTACCGTATAAACCTCAGCAC 3987
QY 541 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCGTTCGAGGACCGAGTCAG 600
Db 3988 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCGTTCGAGGACCGAGTCAG 4047
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGACACTTTTCTTCCACGTC 660
Db 4048 GCAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGACACTTTTCTTCCACGTC 4107
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720
Db 4108 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 4167
QY 721 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 4168 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 4227
QY 781 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 4228 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 4287
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGGCCTGTGAGCTGTGAGCTGATGATGATGATGATGATGAT 900
Db 4288 GAAGGCCCTTTGCTTGGATCTCTTGGCCTGTGAGCTGTGAGCTGATGATGATGATGATGATGAT 4347
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTGAGATTATTAATTTGTTGACCCAC 960
Db 4348 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTGAGATTATTAATTTGTTGACCCAC 4407
QY 961 TATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
Db 4408 TATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 4467
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 4468 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4527
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAG 1140
Db 4528 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAG 4587
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGCCT 1200
Db 4588 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGCCT 4647
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGGGG 1260
Db 4648 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGGGG 4707
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 1301
Db 4708 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 4748

RESULT 6
AX538624
LOCUS Sequence 44 from Patent WO0229056.
DEFINITION AX538624 linear
ACCESSION AX538624
VERSION AX538624.1 GI:25271175
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES
source Location/Qualifiers
1. .11443
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 11443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
DB 6202 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 6261

QY 61 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 6262 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 6321

QY 121 GCCTTTTGAAGGACTAGAGAAACTCTACCCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 6322 GCCTTTTGAAGGACTAGAGAAACTCTACCCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 6381

QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
DB 6382 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 6441

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
DB 6442 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 6501

QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
DB 6502 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 6561

QY 361 CAAGGATCCTGGCAGCCCGGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 420
DB 6562 CAAGGATCCTGGCAGCCCGGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 6621

QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTCGAGCCACGTCAA 480
DB 6622 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTCGAGCCACGTCAA 6681

QY 481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 6682 TGACCTTGCTCGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCAC 6741

QY 541 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGACCGAGTCAG 600
DB 6742 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGACCGAGTCAG 6801

QY 601 GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 660
DB 6802 GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 6861

QY 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
DB 6862 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 6921

QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
DB 6922 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 6981

QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 840
DB 6982 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 7041

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGTGTGATGCCITGGACCCAGCA 900
DB 7042 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGTGTGATGCCITGGACCCAGCA 7101

QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTGACCAC 960

DB 7102 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 7161

QY 961 TATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
DB 7162 TATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 7221

QY 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
DB 7222 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 7281

QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
DB 7282 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 7341

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 1200
DB 7342 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT 7401

QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 7402 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 7461

QY 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301

DB 7462 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 7502

RESULT 7
AX538627
LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1. .12057
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 12057;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCACAGCTTCAGAAGCAGAACGATGTACATAGGCCCTTCAAGAGGGAATTGAA 60
DB 8716 CGACTTTCACAGCTTCAGAAGCAGAACGATGTACATAGGCCCTTCAAGAGGGAATTGAA 8775

QY 61 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 8776 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835

QY 121 GCCTTTGGAAGACTAGAGAACTCTACAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAG 180
DB 8836 GCCTTTGGAAGACTAGAGAACTCTACAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAG 8895

QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
DB 8896 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 8955

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300

8956 AAAATTGAACCTGCACCTCGCTGACTCGGCAGAGAAAAATAGATGAGACCCCTTCAAAGACT 9015
301 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTCAGGTGAT 360
9016 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTCAGGTGAT 9075
361 CAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
9076 CAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 9135
421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
9136 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 9195
481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 9255
541 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
9256 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 9315
601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTC 660
9316 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTC 9375
661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 9435
721 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
9436 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9495
781 TGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCCATGCCATGAAACTCCGAAGACTGCA 840
9496 TGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCCATGCCATGAAACTCCGAAGACTGCA 9555
841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCCTTGGACCAGCA 900
9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCCTTGGACCAGCA 9615
901 CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCCAC 960
9616 CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCCAC 9675
961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 9735
1021 TATGTGCTGAATCGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
9736 TATGTGCTGAATCGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 9795
1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
9796 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 9855
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGCTGGGCT 1200
9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGCTGGGCT 9915
1201 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTGCATCCTTTGGGG 1260
9916 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTGCATCCTTTGGGG 9975
1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
9976 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 10016

LOCUS HSDMDR 12446 bp mRNA linear PRI 12-SEP-1993
DEFINITION Human mRNA for dystrophin.
ACCESSION X14298
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12446)
AUTHORS Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.
TITLE Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str.10, 1115 Berlin Buch, DDR
COMMENT see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES
Location/Qualifiers
1..12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99..11156
/note="unnamed protein product; dystrophin (AA 1 - 3685)";
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="COA:P11532"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVQKKTFTKWNAPQSKFGKHIEFLPSDL QDGRRLDLLEGLTGQKLPKEKSTRVHNLNNKALRVLQNNNVLDVNLGSTDIVDG NHKLTGLIWNIIILHWQKVNKNIMAGLOPTNSEKILLSWVRQSTRNYPQVNLINFT TSWSDGLALNALIHSRPPDLFDWNSVVCQSATORLEHAFNARIYQLGIEKLLDPEDV DTTYDPKKSILMYITSLFQVLPQVSIIEAIOEVEMLPREPVTKEHFQHHQMHYSQ QITVSLAQGYERTSSPKPRFSYAVTQAAVYTTSDTPSPFPSSHLEAPDKSFSGL MESEVNLDRYQTALBEVLWLLSADTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH QGRVGNILQLGSLIGTGLSEDETEVQEQMNLNSWECLRVASMEKQSNLHRVLM DLQNKQKELNDMLTKTEERTKMBEPLGPDLEDLKRQVQHKVLOEDLEQEQVRVN SLTHMVVVDSESGDHATAALEQLKVLGDRWANI CRWTEDRWVLLQDILLKWRILTE EQCLFSAWLSKEDAVNKIHTTGFDQNMELSLVQLEKSTAQISQAVTTQPSLTQTTVM QDLLSTLKNKSVTQTEAWLDNFARCDNLVQLEKSTAQISQAVTTQPSLTQTTVM ETVTVTTRQILVKHAQEEPLPPPPQKKRQITVDSEIRKELDVDTITELHSWITRSEA VLQSPFAIFRKEGNFSDLKEKVNAIEREKAERKLODASRSGQALVEQMNVNIGVNA DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNLQOLEQMTTTAENWLKIQ PTTPSEPTAIKSQLKICKDEVNRLSGIQPIERLKIQSIALKEKGQGPMLDADFVAF TNHFKQVFSVQAREKELQTI FDTLPMPRYQETMSAIRTWVQOSETKLSIPQLSVTDY EIMEQRLGELQALQSSLOEQQSGLYVLTSTTVKEMSKAPSEISRKYQSEFEEIEGRWK KLSQLVHECQKLEEQMNKLRKI QNHITLKKWMAEVDVFLKEWPALGDSEILKKQL KQCRLLVSDIQTIQPSLNSVNEGQKIKNEABPEFASRLLETTELKELNTQWDHMCQQVY ARKEALKGLEKTVSLQKDLSEHWMNTQAEEYLETDEYKTPDELQKAFEEKMRAX BEAQKQEAQVLLTESVNSVIAQAPVAQEAALKKELETTLTNNYQWLCTRLNGCKTILE EWACWHELLSYLEKANKWLNVEVEFKTKTENIPGAAEISEVLDSLENLMRHSNDP NQIRILAQTLTDGGWDELINEELETFNSRWRELHEEA VRRKQKLEQSIQSAQETENS LHIIQESLTFIDKQLAAYIADKVDAQMPQEAQIKQSDLTSHSIELEEMKHNQKGA AQRVLSQIDVAQKQLQDVSMKFRLLFOKPAFQORLOESKMILDEVKMLPALETKSVE QEWQSQLNHCNVNLYKSLSEVKSEVEMVKTGRQIVQKQTPENPKELDERVTALKLHY NELGAKVTRKQLEKCLKLSRMRKEMNVLTETWLAATDMLTKRSAVEGMPNSLDSE VAWGKATQKEIEKQVHLKSI TEVGEALTKVLGKTEIDEDKLTSLNSNWIATVSRAE EWLNLALLEYQKHMETFDQNVDHITKWI IQADTLLESEKKKPOKQEDVLKRLKAELND IRPKVDSTRDQAANLMAHNGDCHKRLVEPQISELNHRFAAISHRITKGKASIPLKELE QFNSDIOQLLEPLEAEIQQGVNLKEEDFNKMDNEDNEGTVKELLQRGDNLQQRITDER

KSEI IKIKOQLLQTKHNAL KDLSRQRKKALEI SHQYQYKROADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELOKKKEELNAVRQAEGLSDEGGAAMAVEPTQIQLSKRWRE
IESKFAQFRRLNFAQIHTVREETMVMTEDEMPLEISYVPSTYLTEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQEESLKNIDSLQOSSGRIDIHSHKKTAAALQSATPVERVKL
QEAALSQDLDFQWEKVNKMYKDRQGRFDRSVKWRFRHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWLXELQDGIQORQTVVRINATGEEIIQQSSKTDA SILQEKLSINL
RWQEVCKQLSDKRKLEEQKNILSEFQDLNEFVLADNFIASIPLEPGKEQOLKE
KLEOVKLLIVEELPLQGIKQLNETGGPVLVSAPISPEEQDKLENKIKOTNLQNTKVS
RALPEKQGEIEAQIKDLGQLEKJEDLEELHLLWLSPIRNLQLEIYNQPNQEGPFD
VKETEIAVQAKQPDVEEILSKGQHLXKEKPAQPVKRLSDLSSEWKAVNRLQILRA
KQPD LAPGLTTIGASPTQTVTLVTPQVVTKETAI SKLEMPSSIMLEV PALADFNRAWT
ELTDWLSLLDQVIKQSRVMVGDLEDINEMIKQKATMQDLEQRRPQLEELITAAQNLK
NKTSNQEARTIITDRIERIQNWDEVQEHQNRRCQNLNEMLKDSTQWLEAKEEABQVL
QCARAKLESWKEGPYVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
DTRKVVHMITENINASWRSIHKRVSEREAALEETHRLLOQFPDLEKFLAWLTEAETTA
NVLQDATRKERLLEDKSGVKELMKQWQDLQGEIEAHTDVYHNLNDSQKILRSLEGSD
DAVLLQRRLDNNMFWSELKKSINIRSHLEASSDQWKRHLSLQJELLVWLKDDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKXKEPVMJMSLTETVRIFLTQEPLEGLEKLYQ
EPRELPPPEERQNVTRLRKQAEVNTWEKLNHLSADWQRKIDETLERLQELQEAETD
ELDLKLRQAEVIKGSQPVGDLLIDSLQDHLKVKALRGEIAPLKENVSHVNDLARQL
TTLGILQSPYNLSTLEDNTRWKLLQVAVEDRVQLHEARDPFGPASQHFSTSVQGP
WERAI SPNKVPYYINHEQTTCWDHPKMTLYQSLADLNNVRSAYRTAMKLLRLQKA
LCDLLLSLSAACDALDQHNLKQNDQPMDIQIINCLTTIYDRLEQEHNNLVNPLCVD
MCLNWLINVDYDTRGTRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCQDQRL
GILLHDSIQIPROLGEVASFGSNIEPSVRSFQFANNKPEIEAALFDWMRLPQSM
VWLPVLRVAAETAKHQAACNICEKPIIGFRYRSIKHFNVDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRKTRKYFAKHPRMGYLPVQTVLEGDNMET
PVTLINFWPVD SAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDISIPNESID
DEHLIIQHYCOSLNODSPLSQPSPAQIILISLEEEERGELERILADLEERNLQAEY
DRLKQOHEHKGLSPLSPPEMPTSPQSPRDAELIAEAKLLRHKGRLEARMQOILEDH
NKQLESQHLRLQLEQPAEAKVNGTTVSSPSTLSQSDSSQPMLLRVVGSQTSDSM
GEBDLSPPQDSTSTGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM"

ORIGIN

Query Match		100.0%;	Score 1301;	DB 9;	Length 12446;		
Best local Similarity		100.0%;	Pred. No. 0;				
Matches 1301;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	CGACTTTCCAGCAGTT	CAGAAGCAGACGAT	GTACATAGGCGCTT	CAGAAGGGAATTGAA	60	
DB	8606	CGACTTTCCAGCAGTT	CAGAAGCAGACGAT	GTACATAGGCGCTT	CAGAAGGGAATTGAA	8665	
QY	61	AACTAAAGBACCTGT	AATCATGAGTACTT	CTTGAGACTGTACGAATAT	TTTCTGACAGAGCA	120	
DB	8666	AACTAAAGBACCTGT	AATCATGAGTACTT	CTTGAGACTGTACGAATAT	TTTCTGACAGAGCA	8725	
QY	121	GCCTTTGGAAGGACT	AGAGAACTCTAC	CAGGAGCCCGCAGAGAGCT	GCCTCTCTGAGGAGAG	180	
DB	8726	GCCTTTGGAAGGACT	AGAGAACTCTAC	CAGGAGCCCGCAGAGAGCT	GCCTCTCTGAGGAGAG	8785	
QY	181	AGCCCAGAAATGT	CACTCGGCTTCT	ACGAAAGCAGGCTGAGGAGGT	CAATACATGAGTGGA	240	
DB	8786	AGCCCAGAAATGT	CACTCGGCTTCT	ACGAAAGCAGGCTGAGGAGGT	CAATACATGAGTGGA	8845	
QY	241	AAAATTGAACCTG	CCTCCGCTGACT	GGCAGAGAAAAATAGAT	GAGACCCCTTGAAAGACT	300	
DB	8846	AAAATTGAACCTG	CCTCCGCTGACT	GGCAGAGAAAAATAGAT	GAGACCCCTTGAAAGACT	8905	
QY	301	CCAGGAACCTTCA	AGAGGCCACGGAT	GAGCTGGACCTCA	AGCTCGGCCAAGCTGAGGTGAT	360	
DB	8906	CCAGGAACCTTCA	AGAGGCCACGGAT	GAGCTGGACCTCA	AGCTCGGCCAAGCTGAGGTGAT	8965	
QY	361	CAAGGGATCTTGC	GAGCCCCCGTGGGCGAT	CTCTCTCATTTGACT	CTCTCTCCAAGATCACCTCGA	420	
DB	8966	CAAGGGATCTTGC	GAGCCCCCGTGGGCGAT	CTCTCTCATTTGACT	CTCTCTCCAAGATCACCTCGA	9025	
QY	421	GAAAGTCAAGGCA	CTTCGAGGAGAAAT	TGCGCCTCTCTGAAAGAGAAC	CGTGAGCCACGTCAA	480	
DB	9026	GAAAGTCAAGGCA	CTTCGAGGAGAAAT	TGCGCCTCTCTGAAAGAGAAC	CGTGAGCCACGTCAA	9085	
QY	481	TGACCTTGCTCGC	CAGCTTACCACTTT	TGGGCATTACAGCT	CTCACCGTATACCTCAGCAC	540	
DB	9086	TGACCTTGCTCGC	CAGCTTACCACTTT	TGGGCATTACAGCT	CTCACCGTATACCTCAGCAC	9145	

QY	541	TCTGGAAGACCTGA	ACACACAGATGGA	AGCTTCTG	CAGGTGGCCGT	CGAGGACCGAGTCAG	600
DB	9146	TCTGGAAGACCTGA	ACACACAGATGGA	AGCTTCTG	CAGGTGGCCGT	CGAGGACCGAGTCAG	9205
QY	601	GCAGCTGCATGA	AGCCCCACAGGAC	TTTGGTCC	CAGCATCTCAG	CACCTTTCTTCCACGTC	660
DB	9206	GCAGCTGCATGA	AGCCCCACAGGAC	TTTGGTCC	CAGCATCTCAG	CACCTTTCTTCCACGTC	9265
QY	661	TGTCCAGGGTCC	TGGGAGAGAGCC	ATCTCGCC	AAACAAAGTG	CCCTACTATATCAACCA	720
DB	9366	TGTCCAGGGTCC	TGGGAGAGAGCC	ATCTCGCC	AAACAAAGTG	CCCTACTATATCAACCA	9325
QY	721	CGAGACTCAAA	CAAACTTGTCTGG	GACCATCCCA	AAATGACAGAG	CTCTACCACTCTTTAGC	780
DB	9326	CGAGACTCAAA	CAAACTTGTCTGG	GACCATCCCA	AAATGACAGAG	CTCTACCACTCTTTAGC	9385
QY	781	TGACCTGAATA	ATGTCTCAGAT	TTCTAGG	ACTTATAGG	ACTGCAACTCCGAAGACTGCA	840
DB	9386	TGACCTGAATA	ATGTCTCAGAT	TTCTAGG	ACTTATAGG	ACTGCAACTCCGAAGACTGCA	9445
QY	841	GAAGGCCCTT	TGCTTGGATCT	CTTGTAG	CCCTGTG	ATGATGCTTGGACCAAGCA	900
DB	9446	GAAGGCCCTT	TGCTTGGATCT	CTTGTAG	CCCTGTG	ATGATGCTTGGACCAAGCA	9505
QY	901	CAACCTCAAG	CAAAATGAC	CCCATG	GATATCCTG	CAGATATTATTAATTTGTTGACAC	960
DB	9506	CAACCTCAAG	CAAAATGAC	CCCATG	GATATCCTG	CAGATATTATTAATTTGTTGACAC	9565
QY	961	TATTTATGAC	CGCCTGGAG	CAAGACACA	CAATTTGGT	CAACGTCCTCTCTGCTGGA	1020
DB	9566	TATTTATGAC	CGCCTGGAG	CAAGACACA	CAATTTGGT	CAACGTCCTCTCTGCTGGA	9625
QY	1021	TATGTGCTGA	ACTGGCTGCT	GAATGTTAT	GATACGGG	ACGAACAGGAGGATCCGTGT	1080
DB	9626	TATGTGCTGA	ACTGGCTGCT	GAATGTTAT	GATACGGG	ACGAACAGGAGGATCCGTGT	9685
QY	1081	CCTGTCTTTT	AAAACTGGCAT	CAITTCCT	TGTGTAA	AGCACATTTGGAAGACAACTACAG	1140
DB	9686	CCTGTCTTTT	AAAACTGGCAT	CAITTCCT	TGTGTAA	AGCACATTTGGAAGACAACTACAG	9745
QY	1141	ATACCTTTT	CAAGCAAGTGG	CAAGTTCA	ACAGGATTTT	GTGACCGCAGCGCTGGCCT	1200
DB	9746	ATACCTTTT	CAAGCAAGTGG	CAAGTTCA	ACAGGATTTT	GTGACCGCAGCGCTGGCCT	9805
QY	1201	CCTTCTGCA	TGATCTATCC	AAATTC	CAAGACAGT	TTGGTGAAGTTGCATCCTTTGGGG	1260
DB	9806	CCTTCTGCA	TGATCTATCC	AAATTC	CAAGACAGT	TTGGTGAAGTTGCATCCTTTGGGG	9865
QY	1261	CAGTAACAT	TGAGCCAA	AGTGTCCG	GAGCTTCC	CAATTG 1301	
DB	9866	CAGTAACAT	TGAGCCAA	AGTGTCCG	GAGCTTCC	CAATTG 9906	

RESULT 9
AX409637
LOCUS
DEFINITION
Sequence 2284 from Patent WO0229103.
AX409637
ACCESSION
AX409637.1 GI:21442342
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
1
AUTHORS
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE
Gene expression profiles in liver cancer
JOURNAL
Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1. .13957
/organism="Homo sapiens"

AX409637
Sequence 2284 from Patent WO0229103.
AX409637
AX409637.1 GI:21442342
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
Location/Qualifiers
1. .13957
/organism="Homo sapiens"

ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 8775

61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835

121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 8836 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCCAGAGAGTGCCTCCTGAGGAGAG 8895

181 AGCCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 8896 AGCCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 8955

241 AAAATTGAACCTGCCTCCGCTGACTGGCGAGAAAAATAGATGAGACCCCTTTGAAAGACT 300
Db 8956 AAAATTGAACCTGCCTCCGCTGACTGGCGAGAAAAATAGATGAGACCCCTTTGAAAGACT 9015

301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 360
Db 9016 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 9075

361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
Db 9076 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 9135

421 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 480
Db 9136 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 9195

481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATACCTCAGCAC 540
Db 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATACCTCAGCAC 9255

541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 9256 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 9315

601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 660
Db 9316 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 9375

661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435

721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 9436 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9495

781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db 9496 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 9555

841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 900
Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 9615

901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTTGTTGACCAC 960
Db 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTTGTTGACCAC 9675

961 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
Db 9676 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 9735

1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 9736 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 9795

1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 9796 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 9855

1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGAGGCTGGGCCT 1200
Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGAGGCTGGGCCT 9915

1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 9916 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975

1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 9976 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 10016

RESULT 10
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1. 13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 8775

61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835

121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCCAGAGAGTGCCTCCTGAGGAGAG 8895

181 AGCCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 8896 AGCCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 8955

241 AAAATTGAACCTGCCTCCGCTGACTGGCGAGAAAAATAGATGAGACCCCTTTGAAAGACT 300
Db 8956 AAAATTGAACCTGCCTCCGCTGACTGGCGAGAAAAATAGATGAGACCCCTTTGAAAGACT 9015

301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 360
Db 9016 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 9075

361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
Db 9076 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 9135

421 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 480
Db 9136 GAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 9195

481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATACCTCAGCAC 540
Db 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATACCTCAGCAC 9255

541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 9256 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 9315

601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 660
Db 9316 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 9375

661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435

721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 9436 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9495

781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db 9496 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 9555

841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 900
Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 9615

901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTTGTTGACCAC 960
Db 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTTGTTGACCAC 9675

||||| 9016 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGAT 9075

QY 361 CAAGGGATCCTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420

Db 9076 CAAGGGATCCTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 9135

QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 480

Db 9136 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 9195

QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCAC 540

Db 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCAC 9255

QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600

Db 9256 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 9315

QY 601 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660

Db 9316 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 9375

QY 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

Db 9376 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435

QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCAAAATGACAGAGCTCTACCACTCTTTAGC 780

Db 9436 CGAGACTCAAAACAACCTTGCTGGGACCATCCAAAATGACAGAGCTCTACCACTCTTTAGC 9495

QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840

Db 9496 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCAGCA 900

Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCAGCA 9615

QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGTCAGATTATTAATTGTTGACCAC 960

Db 9616 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGTCAGATTATTAATTGTTGACCAC 9675

QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTTGGTCAACAGTCCCTCTCTGCGTGGG 1020

Db 9676 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTTGGTCAACAGTCCCTCTCTGCGTGGG 9735

QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 1080

Db 9736 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 9795

QY 1081 CCTGTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAG 1140

Db 9796 CCTGTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAG 9855

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACCGCAGCGTGGGCCT 1200

Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACCGCAGCGTGGGCCT 9915

QY 1201 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260

Db 9916 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975

QY 1261 CAGTAACATTGAGCCCAAGTGTCGGGAGCTGCTTCCAATTG 1301

Db 9976 CAGTAACATTGAGCCCAAGTGTCGGGAGCTGCTTCCAATTG 10016

RESULT 11 HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000

LOCUS Homo sapiens dystrophin (DMD) mRNA, complete cds.

DEFINITION M18533 M17154 M18026 M20250

ACCESSION

VERSION M18533.1 GI:181856

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1699) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.

TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals

JOURNAL Cell 50 (3), 509-517 (1987)

MEDLINE 87273512

PUBMED 3607877

REFERENCE 2 (bases 1678 to 3830) Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.

AUTHORS Conservation of the Duchenne muscular dystrophy gene in mice and humans

TITLE Science 238 (4825), 347-350 (1987)

JOURNAL 89018015

MEDLINE 3659917

PUBMED

REFERENCE 3 (bases 1 to 13957) Koenig,M., Monaco,A.P. and Kunkel,L.M.

AUTHORS The complete sequence of dystrophin predicts a rod-shaped

TITLE cytoskeletal protein

JOURNAL Cell 53 (2), 219-226 (1988)

MEDLINE 88194521

PUBMED 3282674

COMMENT On May 25, 2000 this sequence version replaced gi:340693. Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.

FEATURES

Location/Qualifiers

1. .13957

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xp21.3-p21.1"

/tissue_type="muscle"

/dev_stage="fetus"

/note="G00-119-850"

1. .13957

/gene="DMD"

209. .11266

/gene="DMD"

/codon_start=1

/product="dystrophin"

/protein_id="AAA53189.1"

/db_xref="GI:181857"

/translation="MLWBEVEDCYEREDVQKFTTKWNAQFSKFGKQHIENLFSDDL QGRRLLDLLEGLTGQKLPEKGSTRVHALNNVKNALRVLQNNNVLDLVNIGSTDIVDG NHKLTGLIWNIIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRYPQVNVINFT TSWDGLALNALIHSRPLFDWNSVCCQSATQSLHAFNIARYQLGIEKLLDPEDV DTTYDPDKSILMYITSLFVLPQQVSIEAIEVEMLPRPFPKVKKEHFQLHHQMHYSQ QITVSLAQGYERTSSPKPRFKSYATQAAVYVTSPTSPSPFSSQHLAPEDKSFSSSL MESEVNLDRYQTALEEVLSWLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH QGRVGNILQGSKLIGTKLSEDEETEVEQEQMNLNLSRWECLRVASMEKQSNLHRVLM DLQNKLELNDWLTKEERTKMEEPGLPDLKQVQKHVLLQEDLEQEQVRVN SLTHMVVVVDES SGDHATAALEEQKVLGDRWANI CRWTEDRWVLLQDILLKWQRLTE EQCLFSAWLSEKEDAVNKIHTGFKDQNEMLSSQLKLAVLKADLEKKQSMGKLYSLK QDLLSTLKNKSVTQKTEAWLDNFARCDWNLVQLEKSTAQISQAVTTTQPSLTQTTVM ETVTVTTRERQILVKHAQELPPPPPKRQITVDSEIRKRLDVIDITELHSWITRSEA VLQSPFAIFRKEGNFSDLKEKVNAIEREKAEKFKLQDASRSQAALVEQMVNEGVNA DSIKQASEQLNSRIEFCQLLSERLNWLEYQNNIIAFYNLQQLQEQMTTFAENWLKIQ PTPPSEPTAKSQLKICKDEVNRLSGLQPIERLKIQSLALKEKQGQPNFLDADFVAF TNHFQVPSDVOAREKELOTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY EIMEQRLGELQALQSSLQEQSGLYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK KLSQLVEHCQKLEEQMKNRKIQNHIIQTLKKWMAEVDVFLKEWPALGSEILKKQL KQCRLLVSDIQTIQPSLNSVNEGGQIKNEAEPFASRLETTELKELNTQWDHMCQQVY

gene

CDS

ARKEALKGGLKTVSLQKDLSEMHWMTOAEBYLDERDFEYKTPDELOKAVEEMKRAK
EEAQOKEAKVKLLTSESVNSVIAQAPVPAQEAALKKELETLTTNYQWLCRLNGKCKTLE
EYWACWHELLSYLEKANWLNVEVFKLKTENIPGGAEEISEVLDLSLENLMRHSNDP
NQIRILAQTLTDGGVMDLINELETFNSRWELHEEAVRRQKLELQESIQAASETEKS
LHLIQESLTFIDKQALAYIADKVDAAQMPQEAQIKQSDLTSHESIEEMKKNQOKEA
AQRVLSQIDVAQKQLQDVSMKFERLFQKPAFELRQESKMILDEVKHLPALETKSVE
QEVVQSLNHCVNLKSLSEVKSEVMVITKGQVQKQNTENPKBELDERVETALKHY
NELGAKVTERKQOQLEKCLKLSRKMRKEMNVLTWLAATDMELTKRSAVEGMPNSLDS
VAMGKATQKEIEKQKHLKSI TEVGEALKTVLGKKTLEVEDKLSLNSNWI AVTSRAE
EWNLALLEYQKHMETPDQNVDHITKWI IQADTLLDESEKKPQOQEDVLKRLKAELND
IRPKVDS TRDQAAANLMNRGDHCKLVEPOISELNRHFAAISHRIKTGKASIPLKELE
QFNSDIQKLEPLEAEIQGVNLEKEEDFNKMNEDNEGTVKELLRQGNLQORITDER
KREBIKIQOQLQTKHNAKLRLRSQRRKKALEISHQWQYKQRADDLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSDEGAAMAVEPTQIQLSKRWRE
IESKFAQFRLNFAQIHTVREETMMVMTEDMPLAISYVPSTYLTETHTVSQALLVEQ
LNAAPDLCAQDEFDFKQEEBLKNIKDSLQSSGRIDI IHSKTALQASATPVERVKL
QEALSQLDFOWEKVNKMYKDRQGRFDRSVEKWRFRPHYDIKIFNWLTEAEQFLRKTIQI
PENWEHAKYKWLKELQDGIQORQTVVRTLNATGEEIIQQSSKTASILQEKLSNL
RWQEVCKQLSDRKKRLEEKNI LSEFORDLNEFVLWLEEAADNIASIPLEPGEQOLKE
KLEQVKLLVEELPLRQGLKQNETGGPVLVSAPISPEEQDKLENKLQTNLQWIKVS
RALPEKQGEIEAQIKOLGOLKLEKLEDLEQLNHLHLLWLSPIRNQLBIYNQNOEGPFD
VQETETAVQAQKQPDVEEILSKGQHLKYEKPATQPVKVRKLEDLSSEWKA VNRLLQELRA
KQDPLAPGLTTIGASPTQVTLVTPQVVTKETATSKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLDDQVIKSRVMVGDLEDINEMIIKQKATMQDLRQRRPQJUEELITAAQNLK
NKTSNQEARITIITDRIERI QNWQDEVOEHLQNRQQLNEMLKDSTOWLEAKEEAEQVL
GQARAKLESWEKGYTVDAIQKKITETKQAKDLRQWQTNVDVANDLAKLLRLDYSAD
DTRKVHMITENINASWRSIHKVSEREAALBEETHRLLOQFPDLLEKFLAWL TEAETTA
NVLDQATRKRELLEDSSKGVKELMKQWDLQGEI TEAHTDVYHNDENSQKILRSLEGS
DAVLQORLDNMNFKWSELKKS LNIRSHLEASDDQWKRHLNLSLOQLLVWLQKDDLE
SRQAPIGGDFPAVQKQNDVHRAPKRELKTEPEVIMSTLETVRI FLTEQPLEGLEKLYQ
EPRELPPBEERAQNVTRLRLKQAEVNTWEKLNHLSADWQKIDETLERLQELOEATD
ELDLKLRQAEVIRKIGSWQPVGDLIDLSQDHLEKVKALRGEIAPLEKENVSHVNDLARQL
TTLGTLQSLPYNLSTLEDLNRWKLLOVAVEDRVRLQHEAHRDFGPASQHFLSITSVQGP
WERATSPNKVPYIYNHETQTCWDHPKMTLEYSLADLNNVRFSAYRTAMKLRRLQKA
LCDLLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNPLCVD
MCLNWLINQVDTGRTRIRVLSFKTGIIISLCAHLEDKRYLFKQVASTGFCDDQRRLL
GLLHDSIQIPROLGEVAFGGSNIEPSVRSCTQFANNKPEIEAALFLDWMRLEPQSM
VWLPVLVHRAAAETAKHOAKNICKECPIIGFRYSLKHFNYDICOSEFFSFRVAKGH
KMHYEMVEYCTPTTSGEDVDFDAKVLKNKFKTKYFAKHPRMGYLPVQTVLEGDNMET
PVTLLNFWPVDSAPASSPOLSHDDTHSRIEHYASRLAEMENSGSYLNDISISPNESID
DEHLLIOHYCOSLNQDPSLQSPSPAQILISLESEERGELEIRILADEEENRNLOAEY
DRLKQOHEHKGLSPLSPPEMPTSPQSPRDAELIAEAKLLRQHKHGLEARMQIILEDH
NKQLESQHLRLRQLLEQPOAEAKVNGTTVSSPSTLSQRSOSSQPMLLRVVGSQTSDSM
GEEDLLSPQDTSGLGEEVMEQLNNSFPSSRGRNTPGKPMREDTM"

ORIGIN

Query Match		100.0%;	Score 1301;	DB 9;	Length 13957;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1301;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGACTTTCAGAGTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA	60		
Db	8716	CGACTTTCAGAGTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA	8775		
QY	61	AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	120		
Db	8776	AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	8835		
QY	121	GCCTTTGGAAGGACTAGAGAACTCTACAGAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG	180		
Db	8836	GCCTTTGGAAGGACTAGAGAACTCTACAGAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG	8895		
QY	181	AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAATACCTGAGTGGGA	240		
Db	8896	AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAATACCTGAGTGGGA	8955		
QY	241	AAAATTGAACCTCGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	300		
Db	8956	AAAATTGAACCTCGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	9015		
QY	301	CCAGGAACCTTCAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT	360		
Db	9016	CCAGGAACCTTCAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT	9075		

RESULT 12
AR220819
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

AR220819
Sequence 60 from patent US 6426186.

AR220819
PAT 26-SEP-2002

13977 bp
DNA

linear

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmutth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES location/Qualifiers
1. 13977
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.2%; Score 1290; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 8716 CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 8775
Db |||||
QY 61 AACTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
Db |||||
QY 8776 AACTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 8835
Db |||||
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db |||||
QY 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 8895
Db |||||
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db |||||
QY 8896 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 8955
Db |||||
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
Db |||||
QY 8956 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 9015
Db |||||
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db |||||
QY 9016 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 9075
Db |||||
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA 420
Db |||||
QY 9076 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGA 9135
Db |||||
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db |||||
QY 9136 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 9195
Db |||||
QY 481 TGACCTTGCTCGCCAGCTTACACATTTGGGCATTGAGCTCTCACCGTATACCTCAGCAC 540
Db |||||
QY 9196 TGACCTTGCTCGCCAGCTTACACATTTGGGCATTGAGCTCTCACCGTATACCTCAGCAC 9255
Db |||||
QY 541 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGAGTGGCCGTCGAGACCGAGTCAG 600
Db |||||
QY 9256 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGAGTGGCCGTCGAGACCGAGTCAG 9315
Db |||||
QY 601 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGATCTCAGACATTTCTTTCCACGTC 660
Db |||||
QY 9316 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGATCTCAGACATTTCTTTCCACGTC 9375
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 9376 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435
Db |||||
QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
Db |||||
QY 9436 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 9495
Db |||||
QY 781 TGACCTGAATAATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
QY 9496 TGACCTGAATAATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555
Db |||||

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCA 900
Db |||||
QY 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCA 9615
Db |||||
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
Db |||||
QY 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 9675
Db |||||
QY 961 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db |||||
QY 9676 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 9735
Db |||||
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db |||||
QY 9736 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 9795
Db |||||
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAG 1140
Db |||||
QY 9796 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAG 9855
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCT-GGGCC 1199
Db |||||
QY 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCC 9915
Db |||||
QY 1200 TCCTTCTGCAATGATTTCTATCCAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGG 1259
Db |||||
QY 9916 TCCTTCTGCAATGATTTCTATCCAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGG 9975
Db |||||
QY 1260 GCAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db |||||
QY 9976 GCAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 10017
Db |||||

RESULT 13
AF070485 13887 bp mRNA linear MAM 09-DEC-1998
LOCUS Canis familiaris dystrophin mRNA, complete cds.
DEFINITION AF070485
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 13887)
AUTHORS Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES
Location/Qualifiers
1. 13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1. 289
290. 11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKHLENLPSDL
QDGRRLDLLEGLTGOKLPKEGSTRVHALNNVNKALRVLQKNVDLVNIGSTDIVDG
NHKTLGLIWNIIILHWQVKNVMKIMAGLQQTNSEKILLSWVROSTRNYPQVNVINFT
TMSDGLALNALIHSHRPDLFDWNSVVCOSATQRLHAFNIAKYOLGIEKLLDPEDV
ATYPDKKSILMYTSLFQVLPQQVSEIAIQEIVEMLPSPQVTRBEHFQIHQMHSQ
QITVSLAQGYERAPSPKPKFSYAYTQAAVTTSDPTRSPLPSQHLETPEDKSPGRS
LTETEANLDSYQTALAEVLSWLLSAEDALQAQGEISNDVEEVKEQFHTHEGYMMDLTS
HQRVGNVLQGSOLIGTKLSEDEETEVEQEMNLLNSRWECLRVASWEKQSNLHKVL
MDLQNLKELNDWLTKEERTRKMEKPIGPDIEDLRQVQCHKVLQEDLEQEQVRV
NSLTHMVVWDESSGDHATAALEEQVLKGLDRWANI CRWTEDEWVLLQDILLKQWRF
EEQLFSAWLSEKEDAVNKIHTTGFQDQSEVLNQLKLAFLKTDLEKKKQTMDKLCSL
NQDLLSALKNTVVAHKMEAWLDNFAQRWMDNLVQKLEKSSAQISQAVTTTQPSLTQTTV

5'UTR
CDS

QY 1262 AGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCCAATTG 1301
|||||
Db 10043 AGTAACATTGAGCCGAGTGTCCAGAGCTGCTTCCAGTTG 10082
|||||
RESULT 14
AX306153 13815 bp DNA linear PAT 11-DEC-2001
LOCUS
DEFINITION Sequence 904 from Patent WO0188188.
ACCESSION AX306153
VERSION AX306153.1 GI:17645441
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1. .13815
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCAGCAGTTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 61
|||||
Db 8708 GATTTCAGCAGTTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 8767
|||||
QY 62 ACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAG 121
|||||
Db 8768 ACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGAAATATTTCTGCAGAGCAG 8827
|||||
QY 122 CTTTGGAGGAGTACAGAAACTCTACAGAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGA 181
|||||
Db 8828 CTTTGGAGGAGTACAGAAACTCTACAGAGGAGCCAGAGAGCTGCCTCCTGAGAGAGA 8887
|||||
QY 182 GCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGAA 241
|||||
Db 8888 GCTCAGAAATGTCACCTCGGCTCCTACGAAAGCAGGCTGAGAGGTCAACGCTGATGGGAC 8947
|||||
QY 242 AAATTGAACCTGACCTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTC 301
|||||
Db 8948 AAATTGAACCTGCGCTCAGCTGATTGGCAGAGAAAATAGATGAGCTCTTGAAGACTC 9007
|||||
QY 302 CAGGAACCTTCAGAGGCGCCACGGATGAGCTGGACCTCAAGTTCGCCCAAGCTGAGGTGATC 361
|||||
Db 9008 CAGGAACCTTCAGAGGCTGCCGATGAACTGGACCTCAAGTTGGCCCAAGCTGAGGTGATC 9067
|||||
QY 362 AAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACTCGAG 421
|||||
Db 9068 AAGGGATCCTGGCAGCCAGTGGGGGATCTCCTCATTTGACTCTCTGCAAGATCACTTGAA 9127
|||||
QY 422 AAAGTCAAGGCACTTCAGAGGAGAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAAT 481
|||||
Db 9128 AAAGTCAAGGCACTTCGGGGAGAAAATTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAAT 9187
|||||
QY 482 GACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACT 541
|||||
Db 9188 GACCTTGCAATCAGCTGACCACACTGGGCATTCAGCTCTCACCTTATAACCTCAGCACT 9247
|||||
QY 542 CTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCCGAGTCAGG 601
|||||
Db 9248 TTGGAAGATCTGAATACCAAGATGGAGGCTTCTACAGGTGGCTGTGGAGGACCCGTGTGAGA 9307
|||||
QY 602 CAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCT 661
|||||

Db 9308 CAGCTGCATGAAGCCACAGGGACTTTGGTCTCTGATCCAGCACTTCTTTCCTCACTCA 9367
|||||
QY 662 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCCAC 721
|||||
Db 9368 GTTCAGGGTCCCTGGGAGAGAGCCATCTCACCAACAAAGTGCCTTACTATATCAACCCAC 9427
|||||
QY 722 GAGACTCAAAACAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCT 781
|||||
Db 9428 GAGACCCAAACCACTTTGTTGGACCCACCCCAAAATGACAGAGCTCTACCAGTCTTTAGCT 9487
|||||
QY 782 GACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAG 841
|||||
Db 9488 GACCTGAATAATGTTCAGATTCTCCGCGTATAGGACTGCCATGAACTCAGAAGGCTCCAG 9547
|||||
QY 842 AAGGCCCTTTGCTTGGATCTCTTGGACCTGTGCAGCTGCATGTGATGCCCTTGGACCAAGCAG 901
|||||
Db 9548 AAGGCCCTTTGCTTGGATCTCTTGGACCTGTGCAGCTGCATGTGATGCCCTTGGACCAAGCAG 9607
|||||
QY 902 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACT 961
|||||
Db 9608 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATAATTAAGTGTGACTACA 9667
|||||
QY 962 ATTTATGACCCGCTGGAGCAAGAGCAACAATAATTTGGTCAAGTCCCTCTCTGCGTGGAT 1021
|||||
Db 9668 ATTTATGATCGTCTGGAGCAAGAGCAACAATACTGGTCAATGTCCCTCTCTGTGTGGAT 9727
|||||
QY 1022 ATGTGTCTGAACTGGCTGCTGAATGTTTATGATACCGGAGCAAGAGGATCCGTGTC 1081
|||||
Db 9728 ATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 9787
|||||
QY 1082 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGA 1141
|||||
Db 9788 CTGTCTTTTAAACTGGCATCATTTCTGTGTAAAGCACACTTGAAGACAAGTACAGA 9847
|||||
QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGGCTGGGCTC 1201
|||||
Db 9848 TACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAAGCGTAGGTGGTCTT 9907
|||||
QY 1202 CTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC 1261
|||||
Db 9908 CTTCTGCATGATCTATTCAAATCCCAAGACAGTTGGGTGAAGTTGCTTCTTGGGGGC 9967
|||||
QY 1262 AGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCCAATTG 1301
|||||
Db 9968 AGTAACATTGAGCCGAGTGTGAGGAGCTGCTTCCCAATTG 10007
|||||

RESULT 15
AX538582
LOCUS AX538582 13815 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX538582
VERSION AX538582.1 GI:25271088
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 2 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source 1. .13815
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN

Query Match 87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity 92.1%; Pred. No. 0;

Matches 1197;		Conservative	0;	Mismatches	103;	Indels	0;	Gaps	0;
QY	2	GAC	TTCCAGCAG	TTCCAGAACG	ATGTACATAGG	CCCTTCAAGAGG	GAATTG	AAA	61
Db	8708	GAT	TTCCAGCAG	TTCCAGAACG	ATGATATACATAGG	CCCTTCAAGAGG	GAATTG	AAA	8767
QY	62	ACT	AAAGAACCT	GTAAATCATG	AGTACTCTT	GGAGACTGT	ACGAATAT	TTCTGACAG	AGCAG 121
Db	8768	ACT	AAAGAACCT	GTAAATCATG	AGTACTCTT	GGAGACTGT	ACGAATAT	TTCTGACAG	AGCAG 8827
QY	122	CC	TTTGGAAAG	GAAGGACTAG	AGAAACTCT	ACCAGGAG	CCCCAGAG	AGCTGC	CTTCCAGGAGAGA 181
Db	8828	CC	TTTGGAAAG	GAAGGACTAG	AGAAACTCT	ACCAGGAG	CCCCAGAG	AGCTGC	CTTCCAGGAAAGA 8887
QY	182	GCC	AGAAATGT	CACTCGGCTT	CTACGAAAG	CAGGCTGAG	GAGGTCA	TACTAGTGGG	AA 241
Db	8888	GCT	CAGAAATGT	CACTCGGCTT	CTACGAAAG	CAGGCTGAG	GAGGTCA	TACTAGTGGG	AA 8947
QY	242	AA	ATTGAACCT	GTGCACTCCG	CTGACTGG	CAGAGAAA	AAATAGAT	GAGACCC	TTGAAAGACTC 301
Db	8948	AA	ATTGAACCT	GTGCACTCCG	CTGACTGG	CAGAGAAA	AAATAGAT	GAGACCC	TTGAAAGACTC 9007
QY	302	CAG	GAACCTT	CAAGAGG	CCACCGAT	GAGCTGG	ACCTCAAG	CTGCGC	CAAGCTGAGGTGATC 361
Db	9008	CAG	GAACCTT	CAAGAGG	CCACCGAT	GAGCTGG	ACCTCAAG	CTGCGC	CAAGCTGAGGTGATC 9067
QY	362	AAG	GATCCT	GGCAGCC	CGTGGGCG	ATCTCCTCA	TTGACTCT	CTCTCCA	AGATCACCTCGAG 421
Db	9068	AAG	GATCCT	GGCAGCC	CGTGGGCG	ATCTCCTCA	TTGACTCT	CTCTCCA	AGATCACCTTTGAA 9127
QY	422	AA	AGTCAAGG	CACTTCG	AGGAGAAA	ATTGCGCCT	CTGAAA	GAGAAC	CGTGAGCCACGTCAAT 481
Db	9128	AA	AGTCAAGG	CACTTCG	AGGAGAAA	ATTGCAACCT	TTAAAG	GAGAAATGTCA	ATCGTGTCAAT 9187
QY	482	GAC	CTTGCT	CGCCAG	CTTACC	ACTTTTGGG	CATTGAG	CTCTCAC	CGGTATAAACCTCAGCACT 541
Db	9188	GAC	CTTGCA	CATCAG	CTGTGAC	CACTTGGG	CAATTCAG	CTCTCAC	CTTATAACCTCAGCACT 9247
QY	542	CT	GGAAGAC	CTGAAC	CCAGATG	GAAGCTT	CTG	CAGGTGG	CCGTGAGAGCCGAGTCAGG 601
Db	9248	TT	GGAAGAT	CTGAAT	ACCAGAT	GAGGCTT	CT	ACAGGTGG	CTGTGAGAGCCGTTGTCAGA 9307
QY	602	CAG	CTGCAT	GAA	CCACAGG	CACTTTGGT	CCAGCAT	CTCAG	CACTTTCTTTCCACGTC 661
Db	9308	CAG	CTGCAT	GAA	CCACAGG	CACTTTGGT	CCAGCAT	CTCAG	CACTTTCTTTCCACCTCA 9367
QY	662	GT	CCAGGGT	CCCTGGG	AGAGAC	CCATCTCG	CCCAAAG	AGTGCC	CTACTATATCAACCAC 721
Db	9368	GT	CCAGGGT	CCCTGGG	AGAGAC	CCATCTCA	CCCAAAG	AGTGCC	CTACTATATCAACCAC 9427
QY	722	GAG	ACTCAAA	CAACTT	GTGGGAC	CACTCC	CAAAAT	GACAGAG	CTCTACCAGTCTTTAGCT 781
Db	9428	GAG	ACTCAAA	CAACTT	GTGGGAC	CACTCC	CAAAAT	GACAGAG	CTCTACCAGTCTTTAGCT 9487
QY	782	GAC	CTGAATA	ATGT	CAGATTT	CTCAGCTT	ATAG	AGTGC	CAATGAAACTCCGAGACTGCGAG 841
Db	9488	GAC	CTGAATA	ATGT	CAGATTT	CTCAGCTT	ATAG	AGTGC	CAATGAAAGCTCAGAGGCTCCAG 9547
QY	842	AAG	CCCTTT	GTGATCT	CTTTGAG	CCCTGT	CAGCTGC	ATGTGAT	GCCTTGGACCAGCAC 901
Db	9548	AAG	CCCTTT	GTGATCT	CTTTGAG	CCCTGT	CAGCTGC	ATGTGAT	GCCTTGGACCAGCAC 9607
QY	902	AAC	CTCAAG	CAAAAT	GTACAG	CCCATGG	ATATCT	CTGCAG	ATTATTGTTTGACCACT 961
Db	9608	AAC	CTCAAG	CAAAAT	GTACAG	CCCATGG	ATATCT	CTGCAG	ATTATTGTTTGACTACA 9667
QY	962	ATT	TATGAC	CCGCT	GGAGAG	CAACAATT	TGGTCA	ACGTCC	CTCTCTCGGTGGAT 1021
Db	9668	ATT	TATGAT	CGT	CTGGAG	CAAGAG	CAACAAT	CTGGTCA	ATGTCCCTCTCTGTGTGGAT 9727
QY	1022	AT	GTGCTG	AACTGG	CTGCTGA	ATGTTTAT	GTACGGG	ACGAAC	AGGAGGATCCGTGTC 1081
Db	9728	AT	GTGCTC	CAACTGG	CTTCTCA	ATGTTTAT	GTACGGG	ACGAAC	AGGAGGATCCGTGTC 9787

Search completed: April 5, 2004, 02:39:55
Job time : 3394.8 secs

QY	1082	CTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGA	1141
Db	9788	CTGTCTTTTAAAACTGGCATCATTTCTGTGTAAAGCACACTTGGAAAGACAAGTACAGA	9847
QY	1142	TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTC	1201
Db	9848	TACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTGSGTCTT	9907
QY	1202	CTTCTGCATGATTCTATCCAAATTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC	1261
Db	9908	CTTCTGCATGATTCTATTCAAATCCCAAAGACAGTTGGGTGAAGTTGCTTCCCTTTGGGGGC	9967
QY	1262	AGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG	1301
Db	9968	AGTAACATTGAGCCGAGTGTCCAGAGCTGCTTCCAAATTG	10007